The Impact of State Early Childhood Programs and Child Protective Services Policies on Resilience Following Experiences of Child Maltreatment

by

Sandra Y. Nay McCourt

Department of Psychology and Neuroscience
Duke University

Date: _____

Approved:

_____
Kenneth A. Dodge, Supervisor

_____
Rick H. Hoyle

_____
Timothy J. Strauman

_____
Jacob L. Vigdor

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Psychology and Neuroscience
in the Graduate School of Duke University

2013

ABSTRACT

The Impact of State Early Childhood Programs and Child Protective Services Policies

on Resilience Following Experiences of Child Maltreatment

by

Sandra Y. Nay McCourt

Department of Psychology and Neuroscience
Duke University

Date: _____

Approved:

_____
Kenneth A. Dodge, Supervisor

_____
Rick H. Hoyle

_____
Timothy J. Strauman

_____
Jacob L. Vigdor

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology and Neuroscience in the Graduate School of Duke University

2013

Abstract

In the largest known investigation to date of the prevalence of resilience following experiences of child maltreatment, a statewide, longitudinal sample of maltreated children was used to measure the prevalence of resilience, defined in this study as consistent competence over time and across multiple domains of functioning within the academic setting. In response to the relative paucity of resilience research using large samples, multiple domains of functioning, and longitudinal data, the current study measured resilience in a sample of over 150,000 children who were reported to child protective services agencies for suspected maltreatment. Functioning was measured within three distinct domains (academic performance, special education, and behavioral functioning) across a time period of up to 7 years. A sample of over 450,000 children with no known maltreatment history was used to compare relative rates of consistent competence over time and examine any differential effects on competence across groups. Approximately 18% of maltreated children exhibited consistently competent functioning in all domains across all available years of data, whereas approximately 35% of nonmaltreated children demonstrated consistent competence. County-level introduction of differential response policies investigating children's reported maltreatment was found to promote higher rates of competent functioning. In addition, relative levels of government expenditures in children's counties on two popular statewide early childhood programs (Smart Start and More at Four) were found to predict competent functioning for maltreated and nonmaltreated children alike. These findings suggest that child welfare policies aimed at identifying and assisting high-risk families in need of services and support and community programs targeted at improving children's early development and school readiness hold promise for improving adaptive functioning among maltreated children at high risk for experiencing difficulties in the school environment.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to acknowledge that the current work would not have been possible without the help and support of many individuals and organizations. First, I would like to acknowledge and thank my husband, Tim, for his support, encouragement, and many hours of childcare as I worked on this manuscript. I also thank my children, Madeleine and Carson, for brightening my long work days, evenings, and weekends with their sweet faces and laughter.

I am exceedingly grateful to my mentor and advisor, Dr. Kenneth Dodge, who has provided guidance and support throughout my graduate training, including the process of completing this work. His contributions have been invaluable from the early stages of conceptualizing the current work and planning the study design to the final stages of completing this manuscript. I also appreciate the time and effort of my entire dissertation committee (Dr. Kenneth Dodge, Dr. Rick Hoyle, Dr. Jacob Vigdor, and Dr. Timothy Strauman) in reading and providing feedback on my dissertation proposal and final draft.

I owe a deep debt of gratitude to the North Carolina Department of Social Services, the North Carolina Division of Public Health (Vital Statistics), and the North Carolina Department of Public Instruction for collecting and granting me access to the administrative data used in the current work. I also greatly appreciate the investigators and staff of the North Carolina Education Research Data Center (NCERDC) at the Center for Child and Family Policy, including Dr. Kenneth Dodge, Dr. Clara Muschkin, Ms. Kara Bonneau, and Ms. Sharon Eatmon, for their invaluable work managing and organizing for research use the longitudinal state education and birth records data that are central to the current work. In addition, I would like to thank the NCERDC staff for their specific efforts on my behalf, including assisting me with planning data requests and preparing my IRB protocol, granting me access to the data housed at the NCERDC, and matching administrative data across multiple sources for the current study.

The current work also was made possible by funding for the author from several sources, including the Department of Psychology and Neuroscience and the Graduate School at Duke University, the Center for Developmental Science at the University of North Carolina, the American Psychological Foundation, and the Doris Duke Charitable Foundation and Chapin Hall at the University of Chicago.

I sincerely appreciate all the assistance that I have received during my graduate education and the dissertation process and hope that I can honor the generous gifts of guidance and support I have received through my future work as a scientist and mentor.

1. Introduction

> Because the vast majority of children are adversely affected by their maltreatment experiences, child abuse and neglect may represent the greatest failure of the caregiving environment to provide opportunities for normal development (Cicchetti & Blender, 2006, p. 249).

Child maltreatment is a devastating life experience that has far-reaching impacts on the development and well-being of its victims as well as the welfare of society as a whole. Children who have been abused or neglected are at substantial risk for maladaptive functioning across a broad range of domains, including internalizing problems such as anxiety and depression, externalizing problems such as aggression and delinquency, poor academic performance, and difficulties with peers and other relationships (Arnow, 2004; Cicchetti & Valentino, 2006; Dodge, Bates, & Pettit, 1990; Kaplan, Pelcovitz, & Labruna, 1999; Putnam, 2003; Widom, 1999). Maltreated children are more likely to engage in risky health behaviors such as substance use, experience teen pregnancy, and have abusive and unstable romantic relationships in adolescence and adulthood (Colman & Widom, 2004; Jackson & Martin, 1998; Thornberry, Ireland, & Smith, 2001; Wilson & Widom, 2008).

Maltreated children who grow up to become troubled or violent adults pose substantial costs to themselves and their communities, through higher health care expenses, greater dependence on public assistance, unemployment and occupational difficulties, reduced productivity and unrealized human capital, and costs associated with crime and delinquency (Arnow, 2004; Stagner & Lansing, 2009; Wang & Holton, 2007). As parents, survivors of childhood maltreatment face the prospect of placing their own children at risk for maladjustment as a result of harmful and ineffective parenting behaviors (Kaufman & Zigler, 1987), extending these personal and societal costs to future generations. Recent analyses of the national economic impact of maltreatment conservatively estimated the annual costs at $103.8 billion (Wang &

Holton, 2007) to $124 billion (Fang, Brown, Florence, & Mercy, 2012), not including incalculable human losses from "the pain, suffering, and reduced quality of life that victims of child abuse and neglect experience" (Wang & Holton, 2007, p. 2).

1.1 Definitions and Incidence of Child Maltreatment

In the United States, definitions of child abuse and neglect are established by each state legislature based on federal standards. The Child Abuse Prevention and Treatment Act (CAPTA) (42 U.S.C.A. §5106g, as amended by the Keeping Children and Families Safe Act of 2003), establishes a minimum standard for defining child abuse and neglect: "Any recent act or failure to act on the part of a parent or caretaker which results in death, serious physical or emotional harm, sexual abuse or exploitation; or an act or failure to act which presents an imminent risk of serious harm." The four major categories of maltreatment recognized under most state laws include neglect, physical abuse, psychological abuse, and sexual abuse, and multiple categories can be involved in a child maltreatment investigation.

In 2011, state child protective services agencies (CPS) in the United States received referrals for more than 6.2 million children and investigated or assessed allegations of physical, emotional, and sexual abuse and neglect for more than 3.0 million children, of which an estimated 681,000 children were substantiated as victims of maltreatment, yielding a rate of 9.1 per 1,000 children (U.S. Department of Health and Human Services, 2012). An estimated 1,570 children died as a result of maltreatment, 82% of them infants, toddlers, and preschoolers age 3 and under (U.S. Department of Health and Human Services, 2012). These figures likely represent conservative estimates of the incidence of maltreatment, as there is significant heterogeneity in unsubstantiated cases, many incidents of child abuse and neglect are believed to go unreported, and child fatalities caused by abuse and neglect are not always identified as such, particularly in infancy and early childhood (Drake, 1996; Ewigman, Kivlahan, & Land, 1993; Kohl, Jonson-

Reid, & Drake, 2009). In addition, children whose CPS cases were unsubstantiated often show similar behavioral and developmental outcomes as those with substantiated cases (Barth et al., 2007; Hussey et al., 2005; Leiter, Myers, & Zingraff, 1994), suggesting that substantiation decisions may not differentiate experiences of maltreatment adequately. Population-based surveys tend to produce much higher estimates of the incidence of maltreatment than administrative records (e.g., Finkelhor, Turner, & Hamby, 2005 [124-138 per 1,000 children]; Straus, 1998 [49-270 per 1,000 children]).

In response to concerns about the limitations of using CPS administrative reports to estimate the incidence of child maltreatment, in 1974 the U.S. Congress mandated the National Incidence Studies (NIS). The NIS periodically collect data using a nationally-representative sample of counties, including reports from CPS as well as community professionals, known as "sentinels," who have contact with children and families through law enforcement agencies, public schools, day care centers, hospitals, social service and mental health agencies, juvenile probation and public health departments, public housing, and community shelters. The NIS estimates include abused and neglected children who were reported to CPS as well as maltreated children who were not reported.

The Keeping Children and Families Safe Act of 2003 (P.L. 108-36) mandated the most recent incidence study, the NIS–4, which collected data in 2005-2006 (Sedlak et al., 2010). The NIS-4 used two standards for assessing maltreatment. The "harm" standard for maltreatment was stringent, requiring an act or omission that resulted in demonstrable harm to the child, and was more restrictive than that generally required for CPS substantiation. The less stringent "endangerment" standard for maltreatment applied to a child who had not yet been harmed if a sentinel thought that the act or omission endangered the child or the maltreatment was substantiated or indicated by a CPS investigation. Using the harm standard, more than 1.25

million children were estimated to have been maltreated in the United States during 2005-2006, or 1 in 58 children, and the more inclusive endangerment standard yielded a substantially higher estimate of nearly 3 million maltreated children, or 1 in 25 children (Sedlak et al., 2010).

The NIS reports consistently have found that CPS did not investigate a majority of maltreated children identified in the incidence studies. Continuing this trend, the NIS–4 concluded that CPS investigated maltreatment in only 32% of cases meeting the NIS standard of harm and 43% of cases meeting the NIS standard of endangerment (Sedlak et al., 2010). For example, in 2005-2006, CPS investigated maltreatment for approximately 3.6 million children and substantiated maltreatment in approximately 905,000 cases (U.S. Department of Health and Human Services, 2008), compared to 1,250,000 children identified as having been "harmed" according to stringent criteria of maltreatment and almost 3 million children identified as having been "endangered" according to more inclusive criteria of maltreatment used by the NIS. These findings lend further support to the contention that reports to CPS and the investigation and substantiation of maltreatment likely underestimate the incidence of child maltreatment. Regardless of the specific measure used, the rate of child maltreatment in this country is alarming and warrants substantial efforts to address this serious public health problem.

1.2 Heterogeneity of Outcomes Among Maltreated Children

In recent years, maltreatment researchers have made considerable progress toward identifying risk factors that contribute to the occurrence of child abuse and neglect and the many adverse outcomes associated with experiences of maltreatment (Belsky, 1993; Cicchetti & Lynch, 1993; Cicchetti & Manly, 2001). In turn, empirical evidence on the causes and consequences of childhood maltreatment has contributed to the development of promising prevention and intervention strategies to reduce the incidence of maltreatment and ameliorate its harmful impacts (Cicchetti & Toth, 2006; Olds et al., 1997). Despite these advances, significant work remains to

be done in devising effective methods for preventing maltreatment as well as improving outcomes for children who have been maltreated (Cicchetti & Valentino, 2006; Toth & Cicchetti, 2006).

One important contribution to the development of effective interventions may be to understand better the heterogeneity of outcomes among maltreated children (Bolger & Patterson, 2003; Cicchetti & Valentino, 2006; Luthar & Cicchetti, 2000). Though many maltreated children later experience significant difficulties across a range of important areas of functioning, some individuals "beat the odds" by demonstrating competence and successful adaptation despite the adversity they faced as children (Flores, Cicchetti, & Rogosch, 2005; Masten & Wright, 1998; McGloin & Widom, 2001). Resilient development involves mastering important developmental tasks in areas such as attachment, emotion and behavior regulation, academic competence, and appropriate social interactions in spite of severe threats to development, such as maltreatment (Luthar, Cicchetti, & Becker, 2000). Resilience does not require exceptional performance within domains; rather, positive adaptation in the context of maltreatment means functioning as well as the average (nonmaltreated) child (Luthar et al., 2000; Masten & Coatesworth, 1998; Robinson, 2000). Greater understanding of resilient responses to maltreatment experiences can inform prevention and intervention efforts by identifying factors and processes that promote resilience among maltreated children and, potentially, elucidating critical points of intervention and mechanisms of change (Bolger & Patterson, 2003; Cicchetti & Valentino, 2006; Luthar & Cicchetti, 2000).

1.3 Theoretical Framework for Study of Resilience

No unitary theoretical approach has emerged for studying resilience in maltreated children; however, the developmental psychopathology perspective appears to be the dominant approach among researchers in this area (Luthar et al., 2000; Masten & Obradovic, 2006).

Developmental psychopathology can be described most simply as "the study of behavioral health and adaptation in a developmental context" (Masten, 2006, p. 47). The developmental psychopathology approach is influenced deeply by developmental systems theory and ecological theories of development (e.g., Bronfenbrenner, 1979, 1986; Lerner, 1998; Sameroff, 2000). The ecological multisystems perspective describes human development in terms of multiple, interconnected ecosystems, wherein changes in one system dynamically interact with and impact other systems. In other words, "children's development is strongly influenced by the family, school, peers, neighborhood, and community contexts in which they live" (Fraser, 1997, p. 4).

Masten (2006) enumerated several core principles that characterize the developmental psychopathology approach to studying human development in context, which provide a useful theoretical framework to guide research on resilience. For example, the mutually informative principle provides an empirical rationale for investigating resilience, asserting that greater knowledge of processes of resilience is important for enhancing understanding of both psychopathology and normative development. This principle also supports the critical role that resilience research, alongside research on normative and deviant development, can play in informing prevention and intervention efforts. The systems principle advocates looking beyond the child's individual characteristics to examine the multiple systems and contexts that also influence his development, including his family, neighborhood, peers, school, community, and public policies, in order to help identify external factors that can enhance or promote resilience processes. The longitudinal principle urges researchers to use longitudinal designs to examine resilience processes over multiple time points in order to better capture the dynamic nature of the construct.

Finally, the developmental principle encourages researchers to take into account important age-relevant tasks at each stage of development when deciding how to operationalize

competence and resilience and also to consider how adversity experienced at different ages may have varying effects on particular domains of competencies. This principle also encompasses the concepts of multifinality, in which children experiencing similar events can travel along distinct developmental pathways, and equifinality, in which children who begin with different life experiences can end up in similar final pathways. The concepts of multifinality and equifinality underlie empirical comparisons in resilience research between children who do and do not exhibit resilience following experiences of maltreatment as well as comparisons between children exhibiting resilience and children with more normative developmental trajectories, who did not experience significant adversity and demonstrate generally expected levels of adaptive functioning. These concepts also can inform the search for factors that promote competence and resilience among children with differing developmental trajectories.

1.4 Previous Studies of Adaptive Functioning in Maltreated Children

Many previous studies of resilience following maltreatment have defined resilience as the absence of clinically-elevated levels of mental health symptoms (Jaffee & Gallop, 2007). Though psychological symptomatology is an important component of children's adaptive functioning, there is growing recognition among researchers that resilience encompasses multiple dimensions, and that maltreated children without significant internalizing and externalizing psychopathology may have difficulties in other important areas of functioning (Kaufman, Cook, Arny, Jones, & Pittinsky, 1994; Luthar et al., 2000). In addition, most studies have assessed resilience at one point in time despite the fact that resilience is not conceptualized as a static condition but rather is intended to reflect adaptive functioning across time. These cross-sectional designs likely overestimate the prevalence of resilience among maltreated children, as many children who have faced significant adversity have been found to manifest competence in some domains and

contexts but exhibit problems in others and demonstrate variable levels of adaptive functioning across time (Jaffee & Gallop, 2007; Luthar et al., 2000).

The limited number of studies that measure resilience in multiple dimensions over time generally have estimated the prevalence of continuous resilience at very low rates. Specifically, Farber and Egeland (1987) and Egeland, Carlson, and Sroufe (1993) reported findings from their prospective study of child maltreatment in a sample of 267 low-income mothers in the last trimester of pregnancy. Maltreatment information was gathered during subsequent home visits and laboratory assessments and from public health records. Resilience among the 44 maltreated children identified in the sample was defined as exhibiting secure attachment at the 12-month and 18-month assessments, manifesting a normative level of autonomous functioning at the 24-month assessment, and exhibiting a normative level of self-awareness and socialization at the 42-month assessment. Resilience at the preschool assessment was defined as demonstrating normative levels of prosocial behavior, compliance, independent functioning, and emotional responsivity.

Approximately 50% of the maltreated children and approximately 67% of the nonmaltreated children in this high-risk sample were deemed competent at the 12-month and 18-month assessments. As toddlers, approximately 40% of the maltreated children and approximately 53% of the nonmaltreated children were competent; at 3.5 years of age, approximately 15% of the maltreated children and approximately 43% of the nonmaltreated children achieved competence; and by the preschool assessment, 20% of the maltreated children and 30% of the nonmaltreated children were deemed competent. None of the maltreated children demonstrated competence in all domains across all assessment periods. The specific number of nonmaltreated children demonstrating consistent competence was not reported, but nonmaltreated children were described as having a much less steep pattern of decline in adaptive functioning over time.

A prospective study of maltreatment by Herrenkohl, Herrenkohl, and Egolf (1994) investigated resilience in a sample of 457 children, including a subset of 249 children who were substantiated as experiencing child abuse or neglect and 208 comparison children. Resilience was assessed at 2 time points – once during elementary school and again in adolescence. In the elementary school assessment, resilience was defined as scoring in the top 40% of the full sample of maltreated and nonmaltreated children on indicators of cognitive/academic, social, and emotional functioning. In adolescence, resilience was defined as continued school attendance or graduation from high school. Twenty-five of the maltreated children (10%) met criteria for resilient functioning in childhood. Twenty-three of these 25 children were reassessed in adolescence and only 14 (61%) were still in school or had graduated from high school. Thus, only approximately 6% of the original maltreated sample met resilience criteria over both time periods. In contrast, approximately 40% of the nonmaltreated children were found to be functioning competently in childhood, and a higher number of nonmaltreated children than maltreated children demonstrated consistent competence across both time periods, though the specific rate was not reported.

Cicchetti and Rogosch (1997) conducted a 3-year longitudinal study of 213 low-income children who were attending a summer camp, 133 of whom had been reported for suspected maltreatment. Resilience was operationalized as high or increased adaptive functioning over 3 consecutive years on indicators of interpersonal behavior (sociability, aggressiveness, and social withdrawal), academic competence (school risk index), and psychopathology (self-reported depression and counselor assessment of internalizing and externalizing behavior). Adaptive functioning was measured relative to the full sample of maltreated and nonmaltreated but at-risk children, rather than adaptive functioning on a normative scale. Maltreated children generally showed low levels of adaptive functioning across the 3 years of assessment. Among the

maltreated children, only about 10% were functioning competently at any one assessment point compared to about 35% of the nonmaltreated group. Many of the maltreated children (40%) displayed consistently low levels of resilience across all 3 years compared to 20% of the nonmaltreated children. Similarly, only 1.5% of maltreated children displayed consistently high levels of resilience compared to 10% of the nonmaltreated children.

Bolger and Patterson (2003) reported findings on resilience in a sample of 107 children with substantiated cases of maltreatment who were assessed annually over 4 years along with a comparison sample of 107 nonmaltreated children matched based on age, gender, ethnicity, school attended, and family socioeconomic status. Resilience was assessed across 4 domains, including internalizing behavior, externalizing behavior, peer acceptance, and academic achievement. The authors used two methods to operationalize resilience. The first method defined resilience as scoring 1 standard deviation above the mean of the full sample (including maltreated and nonmaltreated participants) or higher in at least one domain while not scoring 1 standard deviation below the mean or lower in any other domain. The second approach defined resilience as scoring above the full sample median on a composite measure of resilience. Both approaches measured adaptive functioning in comparison to the full study sample rather than a normative standard. Using the first method, 9 maltreated children (8%) demonstrated competence in at least one domain at any one assessment point while only 1 maltreated child (less than 1%) demonstrated competence in at least one domain across all assessment periods. Using the second approach, 23 maltreated children (21%) demonstrated competence on a composite score of resilience at any assessment point while only 5 maltreated children (5%) demonstrated competence across all 4 years.

Jaffee and Gallop (2007) examined the prevalence of resilience using data from the National Survey of Child and Adolescent Well-Being with a sample of 2,065 children who had

been investigated for maltreatment by child protective services. No comparison group was included in the study. Resilience was defined as having no clinically significant internalizing and externalizing problems, demonstrating social competence, and scoring above the mean on standardized reading and math tests. A total of 13% of the maltreated children were deemed competent in all domains of functioning at Wave 1, 14% were deemed competent at Wave 3, which was 18 months after baseline, and 11% were deemed competent at Wave 4, which was 3 years after baseline. Only 2% of maltreated children were deemed to demonstrate continuous resilience over the 3 years of the study.

In summary, estimated prevalence rates of resilience to date have varied widely within and across studies. In longitudinal designs assessing resilience across multiple domains of functioning, a small to moderate percentage of maltreated children has been found to display competence in important developmental tasks during at least 1 assessment point, with rates varying from approximately 10-20% at any one time point. However, the prevalence of resilience was considerably lower when calculated across all assessment periods, ranging from approximately 0-6%.

One notable limitation of research on the prevalence of resilience to date is the fact that most studies have used fairly modest sample sizes to study relatively rare phenomena. The typical samples of 100-300 participants may be more than adequate to study characteristics of maltreated children or even differences between maltreated and nonmaltreated children. This may be particularly true when using high-risk samples intended to oversample for the experience of maltreatment, which is a low-frequency event in the general population. However, the group of most interest to resilience researchers is those maltreated children who manifest resilience, usually comprising a very small portion of the overall sample, as described in the studies above. Even in a large study of more than 2,000 maltreated children by Jaffee and Gallop (2007), the

group of maltreated children exhibiting continuous resilience over the 3 years of the study represented only 2% of the full sample, which would have yielded a group of approximately 40 children for comparative analyses. There is a clear need for larger samples than those used to date in order to study this focal group in a meaningful way, detect significant differences between maltreated children who do and do not demonstrate resilience, and examine factors that promote resilience in this important but relatively small subsection of the population.

1.5 Community Factors Promoting Adaptive Functioning in Maltreated Children

Given variability in outcomes for maltreated children, it is important to identify factors that contribute to resilience, particularly for policy and intervention design. Most research on adaptive functioning following maltreatment has focused on individual characteristics of children, such as self-esteem, intelligence, perceptions of control, self-regulation, attributional styles, and social information processing abilities, as well as factors operating within families, such as parenting quality and family stability (e.g., Bolger & Patterson, 2001; Cicchetti & Rogosch, 1997; Kim & Cicchetti, 2004; Herrenkohl et al., 1994; Luthar, 1991; see Haskett, Nears, Ward, & McPherson, 2006, for a review). A relatively small number of researchers have investigated the effects of relationships with supportive peers and non-parental adults on resilient outcomes (Bolger, Patterson, & Kupersmidt, 1998; Cicchetti & Rogosch, 1997; Luthar & Zigler, 1991; Schwartz, Dodge, Pettit, Bates, & The Conduct Problems Prevention Research Group, 2000). In contrast, very little attention has been focused on broader community-level factors that may enhance or foster resilience among maltreated children, though these external factors arguably may be most amenable to change through public policy (Haskett et al., 2006). A few studies have examined whether involvement in extra-curricular activities, having a positive school climate, or perceived levels of community social support increases the likelihood of resilience, with mixed

results (Egeland, Sroufe, & Erickson, 1983; Perkins & Jones, 2004; Sagy & Dotan, 2001). Another study found that mothers' perceptions of social control and social cohesion were higher among resilient maltreated children (Jaffee, Caspi, Moffitt, Polo-Tomas, & Taylor, 2007). Overall, research of the effects of community-level factors on resilience has been fairly limited to date.

1.5.1 State Child Protective Services Policies

One community-level factor that may influence adaptive functioning in maltreated children is the quality and nature of CPS response to reports of maltreatment. Though the primary goal of child protection agencies handling maltreatment cases generally is to ensure the immediate safety of the children brought to their attention, the family's involvement with agency staff (and through them, the broader spectrum of relevant community resources) also can provide a critical opportunity to support at-risk children and parents and address a multitude of needs in order to enhance children's developmental outcomes. Parents' access to community resources and services can provide critical relief from the stresses of parenting and may improve parents' psychological functioning and their ability to parent their children safely and competently (Daro & Dodge, 2009).

In recent years, state and local governments have grown increasingly interested in innovative efforts to prevent the recurrence of maltreatment among families involved with CPS. In the mid-1990s, some states began implementing policy initiatives allowing CPS staff to approach maltreatment cases using a differential response model, also known as alternative response. As of 2009, approximately 37 states had experimented with or adopted some type of differential response policy or similar reform (Waldfogel, 2009).

The traditional focus in child protective services has been investigating and trying to substantiate reported incidents of maltreatment and, if warranted, providing emergency services to secure the children's safety. Under a differential response approach, typically only the highest-

risk cases that present the possibility of substantial harm are handled in an investigative track, while lower-risk cases are handled in an alternative track focused on assessing families' specific needs as well as strengths and identifying community services and resources that may help support the family (Merkel-Holguin, Kaplan, & Kwak, 2006). The rationale for differential response is:

> to offer flexibility to tailor the child protection response to the needs and circumstances of the family, to collaborate with families early rather than waiting for serious harm to occur, and to remove faultfinding in order to increase the possibility of parent engagement and, ultimately, child safety (Caplan & Merkel-Holguin, 2008).

In addition, focusing investigative resources exclusively on the most severe cases of maltreatment is believed to allow a more intensive response for these highest-risk families and increase the likelihood of protecting those children most at risk of harm (Loman, 2005).

States have employed a variety of strategies for implementing and evaluating the effectiveness of differential response policies at reducing the recurrence of maltreatment and enhancing families' satisfaction with the child protection system. Though specific programs and levels of evaluation vary widely across states, there is a growing body of evidence that differential response can reduce rates of recurrence of maltreatment, make positive impacts on the provision and timing of services to at-risk families, and improving families' engagement in the assessment process (Waldfogel, 2009).

In one of the most rigorously-evaluated programs, Minnesota introduced its alternative response system in 2001 using an experimental design, in which families in 14 counties who were screened as eligible for an alternative response were randomly assigned to have their cases evaluated under the traditional investigative track or under the alternative response system (Loman & Siegel, 2004b). Evaluations of the Minnesota alternative response experiment found

generally positive results on outcomes of interest. Families whose cases were handled under the alternative response approach were less likely to have a repeated report of maltreatment and received more services, particularly preventive services that provided support for general family welfare, such as employment assistance, housing, and child care (Loman & Siegel, 2004b; Siegel & Loman, 2006).

In addition to examining primary program effects, the evaluators also surveyed families regarding broad indicators of family well-being (including income, stress in the home, job-related stress, substance use, domestic violence, access to health care for their children, and ability to care for their children), with more favorable effects found in families served through alternative response. The evaluation examined parents' reports about their children's problems (such as academic, social, and emotional problems) and found that families assessed under the alternative response approach generally endorsed lower levels of problems and showed a decrease in reported problems over time, while the families in the investigative track reported slightly higher levels and tended to show an increase in reported problems over time (Siegel & Loman, 2006).

At least two states have conducted quasi-experimental studies of differential response policies. Missouri began its reform efforts in 1994, implementing a differential response system in 14 counties and evaluating program effects compared to a control group of counties from similar areas. Extensive evaluations of the program found that the program families experienced lower rates of recurrence of maltreatment, received services earlier, and utilized more community resources (Loman & Siegel, 2004a). Similarly, North Carolina initiated a differential response system in 10 pilot counties in 2002, assessing program effects compared to a matched group of comparison counties. Evaluations of program effects found that families assessed by CPS in the pilot counties received more services than in comparison counties, but initial program evaluations did not find a decrease in the recurrence of maltreatment reports between pilot and comparison

counties (Center for Child and Family Policy, 2004, 2006). However, an important finding was that, across all counties, the likelihood of a repeated report of maltreatment within 6 months was reduced in families who received more services early in the assessment process (Center for Child and Family Policy, 2006). In a subsequent follow-up evaluation of the pilot counties and a second wave of counties implementing differential response, program effects on reducing recurrence of maltreatment emerged (Center for Child and Family Policy, 2009; Lawrence, Rosanbalm & Dodge, 2011).

The evidence to date on the effects of state differential response initiatives indicates that these policies have the potential to improve the quality and nature of assessment and services provided by CPS agencies to maltreated children and their families. The outcomes assessed in evaluations of differential response approaches typically include rates of recurrence of maltreatment reports, access to services, family engagement in the decision-making process, and family and staff satisfaction with the program. While these are important outcomes to examine when evaluating whether a program has satisfied its goals, these evaluation studies do not address whether these policies enhance children's adaptive functioning following the experience of maltreatment. Only one evaluation to date (Siegel & Loman, 2006) appears to have made an attempt to examine longer-term developmental outcomes of children. However, the method used to survey parents produced a somewhat limited sample that was not necessarily representative of the families in the study, and the parents' survey responses also may not accurately reflect their children's problems. It does not appear that any rigorous evaluation to date has assessed whether differential response policies have a longer-term impact on children's educational development and adaptive functioning.

1.5.2 State Early Childhood Programs

Another community-level factor that may influence adaptive functioning in maltreated children is the amount of funding allocated to early childhood programs within communities. Local government expenditures represent a broad measure of society's investment in early childhood, collapsed across a range of programs, including interventions for at-risk families. Early investments in children are believed to set at-risk children on a more adaptive course of functioning despite environmental risks, such as poverty, and prepare them better for important life challenges, including school (Heckman & Masterov, 2007; Heckman, 2006). Community services provided to families and children through early childhood programs have the potential to better prepare children to withstand assaults on their development resulting from maltreatment experiences, when provided prior to the occurrence of maltreatment, or may mitigate maltreatment's detrimental effects, when provided after maltreatment has occurred.

Out of growing concerns over the achievement gap beween low-income children and their more-advantaged peers, federal and state governments have invested significant resources in early childhood programs designed to improve school readiness among disadvantaged children. A wide range of early childhood interventions has been implemented and evaluated using different types of study designs, intervention components, intensity and duration of intervention, and size and scale of intervention groups, and the demonstrated effectiveness of these programs has varied across studies (Barnett, 2011; Currie, 2001).

Several programs have used longitudinal randomized trials to test the effectiveness of early childhood interventions. The Perry Preschool Program in Michigan and the Abecedarian Project in North Carolina were small-scale, carefully-implemented studies that produced impressive effects on both short-term and long-term outcomes for program participants in comparison to controls. The Perry Preschool Project provided 2.5 hours per day of high-quality

preschool to the 58 low-income children in the treatment group (Schweinhart et al., 2005). Participants attended preschool for 30 weeks per year for 2 years from ages 3-5. Families also received weekly home visits during the school year. With respect to academic-related outcomes, participants in the Perry Preschool Program obtained higher achievement scores, had better behavior at school, utilized fewer years of special education, and graduated at higher rates compared to the 65 children in the control group. Program participants also had lower involvement in crime and arrests, lower dependency on social welfare, higher earnings, and better health behaviors than their counterparts in the control group in adulthood. The Abecedarian Project provided year-round, full-day educational child care for 5 days per week for 54 low-income children in the treatment group from approximately 4 months of age until age 5 (Ramey et al., 2000). Compared to the 51 children in the control group, Abecedarian participants demonstrated higher achievement scores, lower likelihood of repeating a grade, lower utilization of special education services, and higher levels of high school graduation, educational attainment and skilled employment in adulthood.

Other randomized trials have been conducted on a much larger scale with respect to federal Head Start (N = 4,667) and Early Head Start (N = 3,001) programs. The Head Start randomized trial provided 1 year of preschool education and related family services to 3- and 4-year-old children in the treatment group (U.S. Department of Health and Human Services, 2010). A randomized trial of the Early Head Start program provided child care, parent-child activities, education for parents and adults, and social services for infants and toddlers assigned to the treatment group (Love et al., 2002). Evaluations at ages 2 and 3 found short-term positive effects on children's cognitive and socio-emotional development, but by age 5, the only treatment effect found was a lower level of aggressive behavior (U.S. Department of Health and Human Services,

2006). A subsequent follow-up evaluation in fifth grade found no remaining program effects of the Early Head Start trial (Vogel et al., 2010).

North Carolina introduced statewide initiatives called Smart Start and More at Four that allocated resources to counties to provide early childhood services in order to better prepare disadvantaged children for school entry. Smart Start provided funding to counties to increase access to quality child care, health care, parenting supports, and other community services for children age 0-5 and their families. Smart Start began in 1993 in 18 demonstration counties, and by 1999, it had been implemented in all 100 North Carolina counties. The More at Four Program, introduced in 2001, provided funding to counties to support high-quality preschool programs for eligible 4-year-olds to improve school readiness skills for the year prior to kindergarten. Evaluations of Smart Start found that participation in Smart Start was associated with improvements in the quality of child care centers receiving Smart Start funding and as well as higher levels of cognitive skills and school-readiness among participating children (Bryant et al., 2002; Bryant et al., 2003). Evaluations of More at Four have found that program participation was associated with quality improvements in preschool programs and increases in cognitive and behavioral measures, including higher third-grade test scores compared to a matched sample of children who did not participate in the program (Peisner-Feinberg & Schaaf, 2010). The greatest differences in third-grade test scores were found with respect to the most economically-disadvantaged participants. Another recent evaluation of both Smart Start and More at Four used a quasi-experimental design to evaluate differences in children's educational outcomes based on variations in the dates on which these programs were adopted in various counties and the levels of funding that were received (Ladd, Muschkin, & Dodge, 2012). Children born in counties receiving state allocations for these programs were found to score significantly higher on third-

grade standardized achievement tests in reading and mathematics and be less likely to receive special education services.

Overall, the existing evidence on the effects of early childhood programs suggests that these programs have the potential to improve the educational preparedness and academic outcomes of young children, though mixed results have been found with respect to the lasting impact of these programs. Evaluations of early childhood programs typically have assessed the overall impacts of the program on important outcomes across participants. While policymakers understandably are interested in identifying programs that produce the largest gains on average for a target population, there also is significant benefit to be derived from understanding which programs work best for certain subgroups, particularly those at highest risk for negative outcomes. For example, evaluation studies generally have not addressed whether these broad-based community programs can enhance or promote the adaptive functioning of children who experience maltreatment, although these children are likely to be at significant risk for academic difficulties and school failure. Results from evaluations of early childhood education programs that have examined differential impacts across risk groups of participants suggest that benefits may be largest for the most disadvantaged children (Currie, 2001; Peisner-Feinberg & Schaaf, 2010), and maltreated children represent a particularly disadvantaged group. It does not appear that any evaluation to date has assessed specifically whether early childhood policies have a positive impact on maltreated children's educational development.

1.6 Contributions of the Current Study

The current study aims to make several contributions to empirical knowledge about the prevalence, patterns, and predictors of resilient response to child maltreatment. First, this study examines the longitudinal prevalence of resilience among a large statewide sample of maltreated children in three important domains, academic performance, special education, and behavioral

functioning at school, compared to competence levels in a comparison sample. Using these prevalence data, this study analyzes the stability of resilience over multiple time points. Despite recognition among researchers that resilience is a dynamic and multidimensional developmental construct, previous research on resilience following maltreatment experiences typically has measured resilience at only one time point or in only one domain of functioning, or both (Jaffee & Gallop, 2007; Walsh, Dawson, & Mattingly, 2010). The few existing longitudinal studies measuring multiple outcomes either use small samples from which it is difficult to generalize, or have a relatively brief follow-up period, limiting the ability to estimate stability of resilience over time. There is a clear need for additional longitudinal studies that can examine long-term resilient outcomes for maltreated children in multiple domains of functioning.

In addition, this study uses a large statewide sample to study the relatively uncommon phenomenon of resilience following the experience of maltreatment. Generating a sample that contains a large number of children who have been maltreated and who manifest resilience provides the opportunity to study patterns and predictors of resilience in this focal group in a more robust manner than could be accomplished with the comparatively modest samples typical of research in this area to date.

Second, the current study evaluates the impact on resilient outcomes of recent child welfare reform in North Carolina – the introduction of the Multiple Response System, beginning in some counties in 2002 and expanding to other counties over several years. The Multiple Response System represented a paradigm shift in state child protective services. A fundamental policy change involved implementing a differential response system for interventions with high-risk families (Center for Child and Family Policy, 2004, 2006, 2009; North Carolina Department of Health and Human Services, 2003). Under this system, reports of maltreatment were placed on either a traditional investigative track, for more severe abuse allegations or cases that might

require court involvement, or a new family assessment track. The family assessment track was designed to help stabilize high-risk families and address a broad range of needs that might interfere with effective parenting.

The rationale underlying inclusion of the family assessment track was to foster respect-based partnerships with families, rather than relying on more authoritarian tactics that might alienate parents and reduce the likelihood of cooperation and service uptake (North Carolina Department of Health and Human Services, 2003). These family partnerships were seen as a means of better serving the needs of at-risk families through eliciting more complete information and participation and "buy-in" from parents while still monitoring the family situation closely to protect children from harm. Caseworkers utilizing the family assessment track focused on building upon the family's existing strengths and support systems while engaging the family in services and resources in the community that could enhance the parents' ability to care for their children safely, thereby reducing risks for future maltreatment. The Multiple Response System also focused on tailoring service provision to the relative needs and strengths of each family, providing more intensive interventions to the highest-risk families and less intensive services to families with fewer needs.

The Multiple Response System was piloted in 10 North Carolina counties in 2002, and after receiving an initial positive response, the system was expanded to 42 additional counties in 2003 and implemented statewide in all 100 counties by 2006 (Center for Child and Family Policy, 2006). Expected changes from the implementation of this family-centered approach to child welfare practice included a more comprehensive array of services being provided earlier and to a wider range of at-risk children and families. Building upon existing social supports and family strengths was intended to help foster a natural network to support the family's needs after CPS involvement concluded and empower families to solve their own problems (Center for Child and

Family Policy, 2004, 2006). The current study tests whether this policy change in agency response to at-risk children and families predicts long-term resilient outcomes among children investigated or assessed for maltreatment after the introduction of the Multiple Response System in their county.

Third, the effect of local government expenditures on young children in promoting resilient outcomes among maltreated children is examined. The current study contributes to empirical knowledge of the critical role that local governments can play in addressing the needs of maltreated children and can inform policy decisions regarding effective levels of expenditures to benefit high-risk children. Specifically, this study tests the impact on resilient outcomes of community resources allocated to early childhood programs during the first five years of life.

The current study focuses on two state policies directed at early childhood programs in North Carolina. In 1993, North Carolina initiated its nationally recognized Smart Start program, which was intended to create a comprehensive system of care and education for all children age 0-5, particularly focusing on improving children's preparation for school. Smart Start's goals included improving the quality of child care and early education programs, providing parental supports to encourage positive parenting practices, and increasing access to preventive health care and developmental screenings for young children. Smart Start began as a demonstration project in 18 select counties and was extended gradually to all 100 North Carolina counties by 1998-1999. State funds were allocated to each participating county, and local governments were given discretion to implement and administer programs designed to enhance delivery of community services and resources to young children and families residing within that community, including child care subsidies.

In 2001, North Carolina implemented the More at Four Pre-Kindergarten Program to supplement Smart Start by increasing opportunities for 4-year-old children at risk of school

failure to attend high-quality preschool programs in both public and private schools. Though More at Four was a targeted policy focused primarily on at-risk children, the program also sought to ensure that preschool programs enrolling eligible children met high standards of quality, potentially benefitting not only program participants but also any other children attending the same preschool programs in the community.

The current study tests whether local government expenditures allocated to the Smart Start and More at Four programs to enhance community-level resources available for young children and their families predict long-term resilience among maltreated children residing in that county, and then compare effects to those for nonmaltreated children. Program effects are assessed at the level of community-wide implementation of the programs, rather than examining effects only on children who directly participated in the programs, which may not be feasible given the variety of types and levels of resources different families may have received. This method also allows the full effects of the programs to be evaluated, including direct effects on participants as well as positive and negative spill-over effects to other children within the county (Ladd et al., 2012). Positive spill-over effects might include, for example, improving the overall quality of child care and preschool centers in the community through higher standards imposed on centers enrolling program participants, or the benefits accruing to all children in a classroom when program participants enter school better prepared to learn. Negative spillover effects might include limited access to high-quality care and education opportunities disproportionately provided to program participants and thus reducing available slots for nonparticipants.

It is important to note that much of the current research on factors in resilience following maltreatment experiences (such as establishment of a warm relationship with another adult, positive peer relations, and success in school) is plagued by the problem of possible selection biases. That is, it might be that children who were destined to manifest resilience following

maltreatment selected themselves into particular experiences, without those experiences playing a causal role. The current research solves that methodological problem by identifying pre-existing resources and policies in the community that might contribute to resilience for the maltreated children in that community. These community-level resources are unlikely to be correlated with any individual child characteristics.

1.7 Research Questions

In summary, the current study aims to address the following research questions, in three parts:

I. Measurement of Resilience

1. What is the prevalence of resilience in the domains of academic performance, special education, and behavioral functioning following experiences of childhood maltreatment?

2. How stable is competence in these domains over time, and what are the patterns of instability?

3. How do these prevalence rates and patterns compare to competence in these domains among children with no documented history of maltreatment?

II. Impact of Child Protective Service Reform on Resilience

4. Does living in a community with a comprehensive system of care at the time maltreatment is reported promote resilience among maltreated children?

III. Impact of Resources for Early Childhood Programs on Resilience

5. Do community expenditures on early childhood programs promote resilience among maltreated children, and does this factor operate similarly for nonmaltreated children?

2. Method

2.1 Description of Study

The current study is a longitudinal investigation of resilience among maltreated children in North Carolina in the domains of academic performance, special education, and behavioral functioning. This study involves analyzing administrative data collected by three state agencies in North Carolina: long-form birth records from the Division of Public Health, data contained in reports of maltreatment from the Department of Social Services, and longitudinal education data from the Department of Public Instruction.

2.1.1 Administrative Data

The first data source was provided by the North Carolina Division of Public Health (Vital Records) and made available through the North Carolina Education Research Data Center (NCERDC), housed at Duke University. The data from Vital Records contain individual records on all births in North Carolina from 1996-2003. These birth records included information regarding the child's date of birth, gender, race/ethnicity, birth weight, marital status of the mother, parents' education level, and county of birth.

The second data source included rich administrative records from the North Carolina Department of Public Instruction made available through the NCERDC. One group of important variables in this dataset included standardized end-of-grade test scores and end-of-course test scores for all cohorts between 2001-2012 (where the year refers to the spring of the academic year). These scores are based on state-mandated tests linked to the statewide curriculum and are used in the accountability systems to which all North Carolina schools are subject. Other information obtained from this dataset included whether or not the child was identified as having special education needs (operationalized as having an Individualized Educational Plan), whether the child was identified by the school as limited English proficient, whether the child was retained

in a grade, and whether the child was reported for behavioral infractions, suspensions, or expulsions. Information regarding the specific schools and counties in which participants attended school each year also was tracked over time.

The third data source was provided by the North Carolina Department of Social Services and contained information on reports of maltreatment received by CPS from 1996-2011. Key variables from this dataset included children reported for maltreatment during the relevant years, the specific dates that CPS began investigating reported incidents of maltreatment, and the specific counties investigating reports of maltreatment.

2.1.2 Participants

Individuals eligible for inclusion in the sample of maltreated children included children who could be matched across all three datasets. In other words, the maltreated sample consisted of all children:

(1) born in a North Carolina county between the years of 1996-2003 and living in the state of North Carolina at the time of their birth;

(2) whose families were investigated or assessed by CPS in North Carolina for reported maltreatment between 1996-2011; and

(3) who attended at least one year of public school in the state between 2001-2012.

A comparison group of children born in North Carolina with no documented maltreatment history also was used in some analyses. The comparison group consisted of all children who met criteria 1 and 3 but not 2 above, that is: born in the state between 1996-2003, living in North Carolina at the time of their birth, and able to be matched across the birth records and education datasets. In other words, the comparison sample consisted of all children appearing in the birth records dataset who did not appear in the maltreatment database and for whom at least one year of North Carolina education data was available in the relevant years.

The sample used in the current analyses was identified through the following procedures. Using personal identifiers such as name, birth date, and social security number, staff members at the NCERDC matched all individuals for whom a North Carolina birth record was available for years 1996 through 2003 to the database of individuals for whom the Department of Public Instruction had created master identification numbers for purposes of tracking students across years. In the event that an individual could be matched reliably according to identifying information across databases, the master identification number from the education database was added to the individual birth record file. The merged data files were stripped of identifiers, and only de-identified data were used in the analyses.

The resulting birth records files provided for analysis included 878,143 individuals. Of these files, 16,521 individuals were reported to have resided out of state at the time of birth and were excluded from the analyses. An additional 249,900 individuals with birth records could not be matched to a master identification number in the state education database (most likely, they left the state after birth or attended private schools). The total number of individuals eligible for inclusion in the sample following these matching procedures was 611,722, which represented 70% of the children born in the state during the relevant years. There was no available information on individuals who could not be matched across the birth records and education databases. Potential explanations for matching failures include the possibility that individuals moved to another state prior to enrolling in school, that individuals enrolled in private education or were homeschooled, or that, despite best efforts to match individuals across databases, there were errors in the matching process.

The NCERDC staff conducted a separate matching procedure with respect to maltreatment records from the Department of Social Services and the state education records. The staff matched all individuals for whom a maltreatment report had been filed in years 1996

through 2011 to the master identification numbers in the state education database using personal identifiers. In the event that an individual could be matched reliably according to identifying information across databases, the master identification number from the education database was added to the individual maltreatment record. The merged data files were stripped of identifiers, and only de-identified data were used in the analyses. Maltreatment records were merged across years according to master identification number, and only the first report of maltreatment was retained for the current analyses.

The master list of identification numbers for the 611,722 children who could be matched across the birth records and education databases was matched to the longitudinal dataset of maltreatment records. A total of 156,298 children out of the 611,722 eligible children could be matched across the maltreatment and education databases by master identification number and met initial inclusion criteria for the maltreated group, suggesting a lifetime prevalence rate of maltreatment of 25.6%. This rate is consistent with other estimates and has perhaps greater validity because of its basis in full population data files. The remaining 455,424 children met initial inclusion criteria for the comparison group.

Some children otherwise eligible for inclusion in the maltreated sample only had available education data prior to or within the same academic year as the date their first reported maltreatment incident was investigated or assessed. These children (n = 2,801) were excluded from the analyses because they did not have any eligible years of education data following their first reported experience of maltreatment. The total sample of maltreated children included in the analyses was 153,497.

Review of education data for children eligible to be included in the current study by birth record found no reported education data for any eligible children prior to the 2003-2004 academic year. In the 2003-2004 academic year, only 56 eligible children had available education data, and

in the 2004-2005 academic year, 40,849 eligible children had available education data. By the following 2005-2006 academic year, 319,990 participants were present in the education data, and subsequent years included more participants, with a high of 572,639 participants present in the 2008-2009 academic year. Given the non-representativeness of the available data in academic years 2004 and 2005 compared to academic years 2006 through 2012, the final analyses excluded education data from 2004 and 2005. As a result of the decision to limit the number of years of education data used in the analyses, 489 children in the comparison group were no longer eligible to be included because they did not have available education data in the relevant years. The total sample of children included in the comparison sample for the analyses was 454,935.

In creating the maltreated sample and comparison sample for these analyses, certain assumptions have been made. With respect to the maltreated sample, an underlying assumption is that this sample is representative of all maltreated children in the state. However, it is important to recognize that this sample was limited to maltreated children who continued to live in the state of their birth following maltreatment experiences and attended public schools in the state. It is possible that these characteristics do not represent the larger population of maltreated children and that there are systematic differences between this sample of maltreated children and other maltreated children whose families moved out of the state or engaged their children in alternative schooling arrangements.

Another assumption inherent in utilizing a comparison sample is that the maltreated and comparison groups are identical except for their experiences of maltreatment. However, there may be meaningful differences between these two groups beyond maltreatment status. The analyses were designed to control for certain demographic variables where appropriate to address some of these concerns. In addition, it has been assumed that the children in the comparison group were not maltreated because they did not appear in the Department of Social Services

database. This may not be the case for at least two reasons. First, errors in agencies' administrative record keeping or data entry or errors in the process of matching participants by personal information could mean that the matching procedure unwittingly placed someone who was maltreated in the comparison group. Also, many instances of maltreatment are believed to go unreported, suggesting it is possible that some proportion of participants in the comparison group may have experienced maltreatment but were not reported to CPS.

2.1.3 Data Regarding State Policies and Programs

Other data incorporated into the study included the annual funding levels by county for Smart Start and More at Four in the form of administrative records provided by these programs. These funding levels were matched to participants according to the county in which they first had available education data. Similarly, publicly-available dates of the introduction of the Multiple Response System in all 100 North Carolina counties were matched to maltreated participants based on the counties in which their maltreatment investigations were initiated.

2.2 Measures

2.2.1 Demographics and Covariates

Individual characteristics of sample children and their families were measured and included as covariates in relevant analyses. Table 1 presents descriptive information with respect to the demographic variables and other covariates used in multilevel analyses for the maltreated group, the comparison group, and the overall sample. Demographic measures included the child's gender, race/ethnicity, birth weight, gestational age, and eligibility for free or reduced price lunch as well as the mother's marital status and parents' educational attainment, all of which variables were derived from information contained in the child's birth record or available education records. Other covariates included mother's age at the time of birth, participant's birth year as an indicator of age (ranging from 1996 to 2003), and the number of years of available education data (ranging

from 1 to 7), to allow these characteristics to be controlled statistically in relevant analyses. These latter 3 variables were treated as continuous variables and were grand mean centered to enhance interpretability of parameter estimates in the multilevel logistic regression models (Hox, 2010). Almost half of the participants had available education data for all 7 years included in the analyses. Information on mother's age at birth was missing for a total of 10 participants, and mother's marital status at the time of participant's birth was missing for 15 participants. Those participants with missing data on demographic and other covariates were excluded from multilevel analyses.

Table 1: Descriptive Information on Demographics and Other Covariates

Covariate	Maltreated group	Comparison group	Overall sample
Gender			
Male	51%	51%	51%
Female	49%	49%	49%
Race/ethnicity			
White	49%	60%	57%
Black	35%	23%	26%
Hispanic	5%	8%	8%
Asian	1%	2%	1%
American Indian	1%	1%	1%
Multiracial	9%	6%	7%
Other	0%	0%	0%
Parents' highest education			
Less than high school degree	31%	13%	17%
High school graduate	43%	30%	33%
Some college education	24%	45%	40%
Some graduate education	2%	12%	10%
Free/reduced price lunch			
Eligible	86%	45%	55%
Not eligible	14%	55%	45%
Socioeconomic status (SES)			
Low SES	68%	32%	41%
Middle SES	23%	25%	24%
High SES	9%	43%	35%
Birth weight			
Normal birth weight (≥ 5.5 lbs.)	89%	93%	92%
Low birth weight (3.3-5.4 lbs.)	9%	6%	7%
Very low birth weight (< 3.3 lbs.)	2%	1%	1%
Gestational age			
Full-term birth (≥ 37 weeks)	87%	90%	89%
Late preterm birth (33-36 weeks)	9%	7%	8%
Early preterm birth (≤ 33 weeks)	4%	3%	3%
Fetal well-being			
Full-term birth and normal birth weight	85%	88%	87%
Preterm birth or low birth weight	15%	12%	13%
Mother's marital status at birth			
Married	43%	71%	64%
Not married	57%	29%	36%
Birth year			
1996	14%	12%	12%
1997	13%	12%	13%
1998	14%	13%	13%
1999	13%	13%	13%
2000	13%	13%	13%
2001	12%	13%	13%
2002	12%	13%	13%
2003	9%	11%	10%
Mean (*SD*) birth year	1999.29 (2.223)	1999.51 (2.228)	1999.450 (2.229)
Mean (*SD*) mother's age at birth	23.847 (5.692)	27.213 (5.900)	26.384 (6.028)
Mean (*SD*) years of education data	5.208 (1.882)	5.849 (1.409)	5.688 (1.567)

Note. SES categories were created as follows: Low SES = parent's highest education of less than high school degree and eligible for free/reduced price lunch; Middle SES = parent's highest education of high school degree and not eligible for free/reduced price lunch, or parent's highest education of some college and eligible for free/reduced price lunch; High SES = parent's highest education of some college education and not eligible for free/reduced price lunch, or parent's highest education of some graduate education.

2.2.1.1 Race/ethnicity

Participants were assigned to categories of race/ethnicity based on parents' reported race/ethnicity in participants' birth records. If both parents were reported as being Black, White, or Hispanic, then participant was classified as the same race/ethnicity as both parents. If both parents were reported as being Asian, American Indian, or Other non-White, or if one parent was reported as being from a different race/ethnicity from the other parent, then participant was classified as Other. In the event that information regarding parents' race/ethnicity was missing in participant's birth record (n = 49), then participant's race/ethnicity as reported consistently by schools across available years of education data was used. The percentages of participants in each of the subcategories classified as Other were deemed to be too small for meaningful comparisons with other race/ethnicity categories (American Indian, n = 7,129, approximately 1.2%; Asian, n = 8,171, approximately 1.3%; Other non-White, n = 65, approximately .01%) and were combined with multiracial participants (parents with different reported races/ethnicities) (n = 39,209, approximately 6.4%) to create the category Other. Supplemental analyses using all categories of race/ethnicity produced substantially similar results to those analyses in which the race/ethnicity categories of Asian, American Indian, Other non-White, and multiracial were combined.

2.2.1.2 Socioeconomic status

In order to incorporate two important and distinct yet correlated aspects of socioeconomic status (SES) and to minimize missingness, a proxy variable for SES was created that combined data regarding family income from participants' educational records and parental educational attainment from participants' birth records and educational records. Data from one indicator was used if data on the other indicator was missing, as described below. A chi-square test and Cramer's V was used to assess the association between parental educational attainment and eligibility for free or reduced price lunch. The Cramer's V statistic measures the strength of the

relationship between 2 categorical variables and ranges from 0 to 1.0, with values closer to zero indicating a weaker relationship. A chi-square test using Cramer's V indicated a moderate to strong association between these variables, $\chi^2(3) = 174{,}862$, $p < .0001$, V = .547.

Schools reported whether or not students were eligible for free or reduced price lunch in each available year of education data, and a dichotomous variable reflecting participants' eligibility for free/reduced price lunch across years was created. In order to address inconsistencies in schools' reporting of eligibility for free/reduced price lunch across years and to attempt to reflect participants' family income over time, the following assumptions were made for participants with different free/reduced price lunch eligibility reported across available years of education data. If participants were reported as eligible for free/reduced price lunch in more than one year, they were deemed to be eligible for free/reduced price lunch for purposes of the analyses. If participants had only two years of available data and were designated as eligible for free/reduced price lunch in one year but reported as not eligible for free/reduced price lunch in the other year, they were deemed not to be eligible for free/reduced price lunch for purposes of the analyses.

Each participant's birth record contained data on mother's and father's highest attained education level at the time of participant's birth. In addition, participant's school reported data on parents' educational attainment in several years of educational data. In order to incorporate multiple sources of data and attempt to reflect the possible accumulation of additional years of education after participant's birth, parental education was calculated as the highest education level attained by either parent as reported in participant's birth record or by participant's school. Categories of parental education included less than a high school degree (less than 12 years of education), a high school degree (12 years of education), some college education (13 to 16 years of education), and some graduate education (17 years or more of education).

Participants were classified as low SES if the highest attained parental education was less than a high school degree. Participants also were classified as low SES if at least one parent attained a high school degree but participants were identified as eligible for free/reduced price lunch consistently across available years of data. Participants were classified as middle SES if at least one parent attained a high school degree and participants were not identified as eligible for free/reduced lunch. Participants also were classified as middle SES if the highest parental education attained was some college education but participants were identified as eligible for free/reduced lunch. Participants were classified as high SES if the highest attained parental education included some college education and participants were not identified as eligible for free/reduced lunch. Participants whose parents attained some graduate education also were classified as high SES.

Participants who were missing data on parental educational attainment (n = 473) were classified as low SES if they were identified as eligible for free/reduced price lunch and middle SES if they were identified as not eligible for free/reduced price lunch. Participants who were missing data on free/reduced price lunch eligibility (n = 24,220) were classified as low SES if the highest parental educational attainment was less than a high school degree, middle SES if the highest parental educational attainment was a high school degree or some college education, and high SES if the highest parental educational attainment was some graduate education. A total of 34 participants were missing data on both SES indicators and were excluded from the multilevel analyses.

2.2.1.3 Fetal well-being

Information about participants' birth weight and gestational age reported on their birth certificates also was included as a covariate in the analyses. Participants' birth weight was calculated as a dichotomous variable: normal birth weight (greater than or equal to 5.5 pounds

(2,500 grams) or low birth weight (below 5.5 pounds (2,500 grams)). Though some participants could be classified as very low birth weight (below 3.3 pounds (1,500 grams)), the percentage of participants falling in this category was very low (n = 8,073, approximately 1.3%) and was not deemed meaningful for comparative analyses. Thus, low birth weight and very low birth weight categories were combined. Similarly, participants' gestational age at birth was calculated as a dichotomous variable: full-term birth (37 weeks or more) or preterm birth (less than 37 weeks). Though some participants could be classified as early preterm birth (less than 33 weeks), the percentage of participants in this category was small (n = 18,077, approximately 3.0%) and was deemed not meaningful for comparative analyses. A chi-square test indicated a moderate to strong association between birth weight and gestational age ($\chi^2(1) = 219{,}633$, $p < .0001$, V = .601).

In order to include these two associated indicators that also have independent effects on neonatal health and to minimize missingness, information on low birth weight and preterm birth was combined into a dichotomous variable of fetal well-being. Participants were classified as having been born full-term with normal birth weight or having been born preterm or with low birth weight. Participants who were missing data on birth weight (n = 85) were classified based on gestational age only, and participants who were missing data on gestational age (n = 419) were classified based on birth weight only. A total of 9 participants were missing data on both fetal well-being indicators and were excluded from the multilevel analyses.

2.2.2 Maltreatment Status

All children born in North Carolina in 1996-2003 who were reported to the Department of Social Services for suspected neglect or abuse in 1996-2011 were eligible for inclusion in the maltreated sample, including both substantiated and unsubstantiated reports. The decision to include reports that were not substantiated was informed by many factors, most compelling

among them the fact that the reasons underlying substantiation decisions are heterogeneous and do not indicate that maltreatment did not occur (Drake, 1996; Drake et al., 2003; Ewigman et al., 1993). The reliability of substantiation decisions has been found to be questionable (Slep & Heyman, 2006), and the substantiation determination alone does not distinguish children who are at risk of future maltreatment or families who are in need of services (Kohl et al., 2009). In fact, children whose CPS cases were unsubstantiated have been found to be at similarly high risk of being reported again, having a substantiated case of maltreatment, and being placed in foster care in the future as those with substantiated cases (Drake et al., 2003; Kohl et al., 2009). In addition, developmental, health, and behavioral outcomes for children with substantiated and unsubstantiated cases have been found to be similar (Barth et al., 2007; Hussey et al., 2005; Leiter et al., 1994). Notably, as a result of the Multiple Response System policy reform beginning in 2002, for families in the family assessment track, North Carolina no longer uses the substantiation designation and instead assesses the family's need for services. Relying exclusively on official reporting, including both substantiated and unsubstantiated cases, is likely a conservative measure of maltreatment, with population-based surveys of maltreatment experiences reporting significantly higher rates of maltreatment than rates based on incidents reported to CPS (e.g., Finkelhor et al., 2005; Straus, 1998).

For all of the foregoing reasons, the current study included all eligible children who came to the attention of CPS for suspected maltreatment in the relevant years. Maltreatment status for the purpose of this study was operationalized as a dichotomous variable indicating whether or not there was a record of any CPS report for suspected maltreatment of the participant between participant's date of birth and December 31, 2011. The average age at which participants were first reported for maltreatment was 4 years old ($M = 4.128$; $SD = 3.444$), with ages ranging from 0

to 15 years of age. For children attending school at the time they were reported for maltreatment, grade levels ranged from preschool to 9th grade.

2.2.3 Competence, Resilience, and Eligible Years of Education Data

Resilience was operationalized as a dichotomous outcome variable indicating consistently competent functioning in academic performance, special education, and behavioral domains over all available years following the experience of maltreatment. Participants' competence was calculated separately within each year of available data. Measures of competent academic performance across elementary, middle, and high school years included end-of-grade test scores in Grades 3-8 and Grade 10, end-of-course test scores in approximately Grades 9-12, and grade retention status. Criteria for competence involved achieving grade-level proficiency in all end-of-grade and end-of-course testing (Level III or Level IV achievement level) and not being retained in a grade that year. Competence in the domain of special education was measured by education records reporting students' English-language proficiency and special education status. Competence in the special education domain was operationalized as not being identified by the school as limited English proficient and not having a designation for special education services or accommodations under an Individualized Education Plan. Measures related to behavioral functioning included school-reported infractions, suspensions, and expulsions, and criteria for competence involved the absence of reported behavioral problems. Criteria for competence were the same for the comparison group and the maltreated group. Table 2 specifies the criteria for competence within each domain.

Table 2: Criteria for Competence Within Specific Domains of Functioning

Academic performance	Special education	Behavioral functioning
Proficient in all required end-of-grade achievement testing (Reading, Mathematics, Science, and Writing) and Proficient in all required end-of-course achievement testing (Algebra I, Algebra II, Biology, Chemistry, Civics & Economics, Geometry, English I, Physical Science, Physics, and U.S. History) and Promoted to next grade	Not provided accommodations or services under an Individualized Education Plan and Not identified as Limited English Proficient	Not reported for behavioral infractions, suspensions, or expulsions

For the maltreated group, competence was calculated using only years of education data after participants were reported for maltreatment. Eligible years of data for maltreated participants were limited to academic years following the date on which participant's CPS investigation or assessment was initiated, using September 1 as the relevant comparison date, and any subsequent years of available education data. If a participant was investigated or assessed for maltreatment before September 1, the school year beginning on or about September 1 of that year was deemed to be the first eligible year of education data. For participants investigated on or after September 1, the subsequent academic year was deemed to be the first eligible year of data.

This criterion was implemented to ensure that all outcomes of interest were measured after maltreatment occurred and was investigated. Participants' academic testing occurred at several time points during the school year, typically at the end of the fall semester (for fall semester end-of-year courses), the middle of the spring semester (for spring semester end-of-year courses and most end-of-grade testing), and the end of the spring semester or summer (for remaining end-of-grade testing as well as promotion and retention decisions). School decisions to implement special education services and make reports of behavioral infractions could have

occurred at any point in the school year, and special education decisions and academic performance were likely to have been based on cumulative performance during the school year. Thus, a conservative criterion was imposed to ensure that outcomes were measured after maltreatment occurred and CPS initiated an investigation or assessment. A total of 63% of the maltreated group (n = 96,184) was reported for maltreatment prior to the beginning of the 2005-2006 academic year, which was the first year of education data used in the analyses, and their data were not affected by this criterion.

2.2.4 School County

For purposes of conducting multilevel analyses, a variable was calculated to reflect the county in which each participant's school district was located. Almost all participants in the sample remained in the same school county throughout all available years of education data. Approximately 10% of the participants were reported to have moved counties at least once across available school years (n = 57,873). For those participants who moved, their school county was assigned based on the county in which participant had the greatest number of available years of education data. In the event that a participant had an equal number of years of education data from two or more counties (n = 5,576), that participant was assigned to the most recent school county. A total of 13 participants had no identified county associated with their education data and were excluded from the multilevel analyses.

2.2.5 Introduction of the Multiple Response System and Family Assessments

The initial investigation date for each maltreated participant's CPS report was compared to the date of implementation of the Multiple Response System in the county in which the CPS report was investigated. A dichotomous variable was calculated for each participant, designating the participant as having been reported either prior or subsequent to the introduction of the Multiple Response System within the county investigating participant's CPS report. The Multiple

Response System was introduced in 3 waves: the 10 pilot counties adopted it as of January 1, 2003; the second wave of 42 counties adopted it as of January 1, 2004; and the remaining 48 counties comprising the third wave adopted it as of January 1, 2006. Data also were provided by the Department of Social Services on the first month in which counties began reporting family assessments, providing evidence of the implementation of an important component of the Multiple Response System within counties. Supplemental analyses examined whether there were any differential effects on resilient functioning for participants reported for maltreatment prior to or subsequent to formal implementation of family assessments within the county investigating or assessing their initial maltreatment report.

2.2.6 County Expenditures on Smart Start and More at Four Programs

Data on total county expenditures on the Smart Start and More at Four programs for infants, toddlers, and young children age 0 to 5 for all 100 North Carolina counties were matched to participants' counties of residence in early childhood. County of residence data were derived from each participant's county of birth as well as the county where participant's first year of available education data was reported. Annual county expenditures per capita were calculated for the first 5 years of each participant's life based on participant's birth month and year, using fractional multipliers for partial years, and summed across time. Children whose county of birth was consistent with the county in which their first year of education data was available were assumed to have lived continuously in the same county. For children who appeared to have moved from their county of birth prior to their first year of available education data, two separate sets of analyses were conducted. One set of analyses included these "movers" assuming that these children remained in their birth county through age 5 and thus received the benefit of county expenditures for Smart Start and More at Four exclusively in their county of birth. Another set of analyses included movers assuming they moved to the second county immediately after birth and

benefited from expenditures on Smart Start and More at Four exclusively in the second county. Although neither assumption is likely to have been correct for all children in the movers subsample, these alternative analyses allowed children who moved among North Carolina counties to be included and permitted the robustness of results across assumptions to be evaluated.

In addition to the analyses using the full sample, alternative analyses evaluated the effect of Smart Start and More at Four expenditures using only the "nonmovers" assuming that they lived continuously in their county of birth. It is possible that some children in the nonmover subsample who were born in one county may have moved subsequently to another county with differing levels or different timing of investment in Smart Start and More at Four and then moved back to their birth county prior to starting school. However, without specific information on where such children resided between their birth and the date they entered school, which was not available in the data for this study, the county in which these children lived during their early childhood years could not be identified precisely. Thus, the child's birth county was deemed to be the best available indicator of county of residence during early childhood for the nonmovers.

2.3 Data Analysis

2.3.1 Prevalence of Consistent Competence and Patterns of Competence

In order to examine patterns of stability and instability in competence over time and within specific domains of functioning, the following set of dichotomous variables was calculated for each year of available academic data: met criteria for competence in the academic performance domain; met criteria for competence in the special education domain; met criteria for competence in the behavioral functioning domain; and met criteria for overall competence in that year. A dichotomous variable indicating consistent competence across all available years of education data also was created for each participant. Participants were included in the analyses as long as they had at least one year of education data in North Carolina between 2006 and 2012, but

some participants were missing one or more years of education data during these years. The existence of missing data is not surprising given the longitudinal nature of the data and the complexity of collecting and matching administrative data for a statewide sample. Participants' consistent competence was evaluated only with respect to years of available data. The number of domains within which participants were consistently competent as well as the prevalence of consistent competence within each specific domain also was examined.

In addition to the foregoing analyses focusing on the prevalence and patterns of competence and examining differences between groups, additional analyses tested the statistical significance of differences in the prevalence of consistent competence between maltreated and nonmaltreated children in the sample using multilevel modeling techniques to take into account the nested nature of the data, with children located within their respective counties of residence. The nested nature of these data may produce biased estimates in statistical techniques relying upon the assumption that all observations are independent (e.g., Hox, 2010), as it is plausible that children residing within the same counties, utilizing the same community resources and agencies, and attending schools in the same district may be more similar to each other than to children in other counties in ways that influence their competence.

The possible non-independence of children living within the same county was addressed by using a two-level hierarchical linear model that included individual participants at Level 1 and individual counties at Level 2. The overall sample size of individuals (N = 608,419) and the total number of the highest-level units (N = 102 counties/local education agencies) were deemed to be adequate to support multilevel analysis (Hox, 2010; Newman & Newman, 2012; Raudenbush & Bryk, 2002). Maltreatment status was included as a Level 1 dichotomous predictor variable in a multilevel logistic regression model to assess whether consistent competence across academic performance, special education, and behavioral functioning domains (a dichotomous outcome

variable) differed significantly for maltreated and nonmaltreated children. Covariates added as Level 1 control variables included the child's birth year and the number of years of available education data as well as several demographic variables: the child's gender, race/ethnicity, and fetal well-being, the mother's marital status and age at the time of the child's birth, and the family's socioeconomic status.

Multilevel logistic regression analyses were performed in SAS® 9.3 (SAS Institute Inc., 2011) using PROC GLIMMIX for generalized linear mixed models. The GLIMMIX procedure extends generalized linear modeling techniques to data with nonnormal distributions, including binary outcomes (Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006). Model parameters were estimated via maximum likelihood based on adaptive Gaussian quadrature specifying numerical integration over 15 quadrature points (METHOD=QUAD) (QPOINTS=15). Binary distribution of the response variable (DIST=BINARY) and the logit link function (LINK=LOGIT) were specified for multilevel logistic regression analysis. Multilevel regression analyses were performed in SAS® 9.3 (SAS Institute Inc., 2011) using PROC MIXED for linear mixed models.

2.3.2 Effects of State Policy Reform in Child Welfare

The analysis of the effects of state policy reform in child welfare was conducted using only the maltreated sample and tested whether being investigated for maltreatment in a county that had adopted the Multiple Response System predicted resilient outcomes. Multilevel logistic regression analyses were conducted to reflect possible non-independence of observations from children nested within counties. These multilevel logistic regression analyses tested whether the introduction of the Multiple Response System in the child's county of residence prior to reported maltreatment (a dichotomous independent variable) predicted consistent competence across academic performance, special education, and behavioral functioning domains (a dichotomous

outcome variable). The effects of the introduction of the Multiple Response System on competence within each domain of functioning and on the percentage of years in which participants were competent also were evaluated. The same set of covariates was used as those included in the multilevel logistic regression analyses predicting competence across both the maltreated and comparison samples, as described above. In order to address multicollinearity among two of the covariates and the predictor of interest likely resulting from the underlying common factor of time, the child's birth year and the number of years of available education data were combined into their cross-product, reflecting the interaction of birth year and available years of education data.

2.3.3 Effects of County-Level Expenditures for Early Childhood Programs

The analysis of the effects of government expenditures on Smart Start and More at Four tested whether average county-level expenditures on these programs during participants' early childhood predicted their long-term competent outcomes. Because children were nested within counties, multilevel logistic regression analyses were used to address the non-independence of observations. These models examined whether levels of per capita expenditures on Smart Start and More at Four in the child's county of residence between birth and age 5 (a categorical independent variable) predicted consistent competence across academic performance, special education, and behavioral functioning domains (a dichotomous outcome variable). The same set of covariates was used as those included in the previous sets of analyses. Additional analyses investigated the separate effects of each of the More at Four and Smart Start programs. The effects of these programs on competence within each domain of functioning and on the percentage of years in which participants were competent also were evaluated. Further analyses examined whether there were independent and additive effects of these early childhood programs

on maltreated children when combined with the introduction of the Multiple Response System within participants' county of residence.

2.3.4 Competence Prior to and Following Experiences of Maltreatment

An additional set of analyses focused on the subsample of maltreated participants for whom years of education data were available prior to the year in which they were reported for maltreatment. The prevalence of consistently competent functioning in both pre-maltreatment years as well as post-maltreatment years was calculated, and descriptive categories of relative functioning over time for these participants were created. Multilevel logistic regression models were used to examine whether county expenditures on early childhood programs impacted pre-maltreatment and post-maltreatment competence in this subgroup of maltreated participants.

3. Results

The results of the current study are presented in five sections. The first section describes the prevalence of consistent competence and patterns of competence across years and within each specific domain for maltreated and nonmaltreated children. The second section presents data with respect to maltreated children on the effect of the timing of introduction of the Multiple Response System on consistent competence following experiences of maltreatment. The third section sets forth the effect of government expenditures for early childhood programs on consistent competence for maltreated and nonmaltreated children. The fourth section reports the independent and additive effects of the timing of the introduction of the Multiple Response System and relative expenditures on early childhood programs. The fifth section describes the prevalence of consistent competence and patterns of competence across pre-maltreatment and post-maltreatment years for the subsample of maltreated children for whom pre-maltreatment data were available.

3.1 Competence Across Years and Domains of Functioning

3.1.1 Prevalence of Consistent Competence and Patterns of Competence

In order to permit ready comparisons to previous estimates of the prevalence of resilience in the literature, percentages reflecting the prevalence of consistent competence for all years of available education data following the first report of maltreatment (for the maltreated sample) and consistent competence for all years of available education data (for the comparison sample) were calculated. Table 3 presents these prevalence data, reporting competence within each year and across all years for the overall sample and separately by maltreatment status as well as gender and race/ethnicity. A total of 18% of participants in the maltreated sample demonstrated consistent competence in all available domains across all available years, with 36% to 74% of the maltreated sample exhibiting competent functioning within any one year. In contrast, 35% of participants in

the nonmaltreated sample demonstrated consistent competence, with 59% to 84% of the nonmaltreated sample exhibiting competent functioning within any one year. The difference between the maltreated and nonmaltreated groups on consistently competent functioning (not controlling for any covariates) was significant ($\chi^2(1) = 15{,}809.008$, $p < .0001$).

Table 3: Competence Across Years by Maltreatment Status, Gender, and Race/Ethnicity

Group	2006	2007	2008	2009	2010	2011	2012	All years
Total competent	261,798	308,591	300,214	357,013	307,528	308,806	296,268	188,777
Total not competent	58,192	91,439	185,741	215,626	257,093	251,256	261,000	419,655
Total present in year	319,990	400,030	485,955	572,639	564,621	560,062	557,268	608,432
% competent	82%	77%	62%	62%	55%	55%	53%	31%
Competence by maltreatment status (25% maltreated)								
Total maltreated group	49,382	56,385	50,559	59,246	51,333	51,673	50,929	27,381
% competent	74%	66%	47%	47%	39%	38%	36%	18%
Total comparison group	212,416	252,206	249,655	297,767	256,195	257,133	245,339	161,396
% competent	84%	80%	66%	67%	59%	61%	59%	35%
Competence by gender (51% male)								
Males - Maltreated	24,646	27,573	23,403	26,507	22,025	21,688	21,410	10,594
% competent	72%	63%	42%	40%	32%	31%	29%	13%
Males - Comparison	105,649	124,112	118,648	140,275	118,405	117,129	111,452	68,704
% competent	82%	77%	61%	62%	54%	54%	53%	30%
Females - Maltreated	24,736	28,812	27,156	32,739	29,308	29,985	29,491	16,787
% competent	76%	69%	52%	53%	45%	45%	43%	22%
Females - Comparison	106,767	128,094	131,007	157,492	137,790	140,004	133,887	92,692
% competent	86%	83%	71%	72%	65%	68%	66%	42%
Competence by race/ethnicity (57% White)								
White - Maltreated	25,030	29,064	26,146	32,307	29,941	29,761	29,936	17,652
% competent	77%	69%	50%	52%	47%	45%	44%	24%
White - Comparison	135,656	161,439	161,150	194,641	175,851	177,729	172,366	122,995
% competent	87%	84%	70%	73%	68%	70%	69%	45%
Black - Maltreated	17,268	18,855	16,194	17,136	13,021	13,642	13,136	5,567
% competent	70%	60%	42%	37%	27%	28%	26%	10%
Black - Comparison	47,256	53,261	47,786	54,312	42,545	43,300	40,281	19,739
% competent	77%	71%	53%	52%	42%	44%	41%	19%
Hispanic - Maltreated	1,680	2,236	2,365	2,911	2,277	2,220	1,958	917
% competent	73%	68%	53%	50%	36%	33%	28%	12%
Hispanic - Comparison	13,156	17,616	20,158	24,203	17,162	15,038	12,532	5,645
% competent	81%	77%	66%	64%	47%	42%	36%	15%
Other - Maltreated	5,404	6,230	5,854	6,892	6,094	6,050	5,899	3,254
% competent	76%	68%	50%	49%	42%	40%	38%	19%
Other - Comparison	16,348	19,890	20,561	24,611	20,637	21,066	20,160	13,017
% competent	85%	81%	67%	67%	58%	61%	60%	35%

Note. Race/ethnicity categories of Asian, American Indian, Other non-White, and multiracial have been combined into category "Other."

Figure 1 depicts the relative percentages of maltreated participants who were consistently competent across time by age category when they were first reported for maltreatment. The largest group was reported for maltreatment between birth and age 2, representing 40% of the sample. Twenty-seven percent (27%) of the sample was reported between the ages 3 and 5, 20% of the sample was reported between the ages of 6 and 8, 10% of the sample was reported between the ages of 9 and 11, and the remaining 3% of the sample was reported between the ages of 12 and 15. Relative percentages of consistent competence increased as a function of the age of participants, ranging from 15% to 25% across age category.

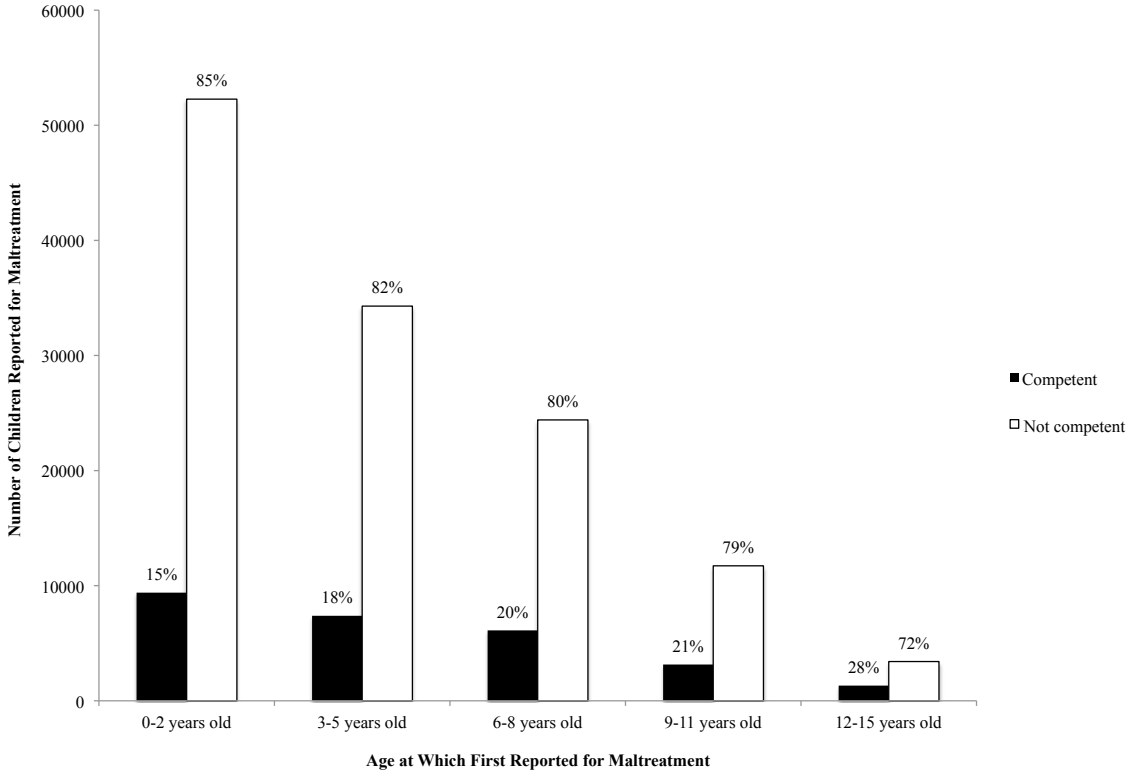

Figure 1: Consistent Competence by Age at Maltreatment

In order to examine patterns of competence across years of education data, the percentage of years in which participants were competent was calculated for the maltreated and nonmaltreated groups. Figure 2 presents the cumulative percentage of participants in each group

who were competent for the stated percentage of years. At the low end, 19% of the maltreated group and 6% of the nonmaltreated group never achieved competent functioning across domains in any year of available education data. However, almost half of the maltreated sample and more than two-thirds of the nonmaltreated sample were functioning competently across all domains in at least 50% of available years. The percentage of years in which the maltreated group and nonmaltreated group were competent (not controlling for any covariates) was significantly different ($z = -197.279$, $p < .0001$).

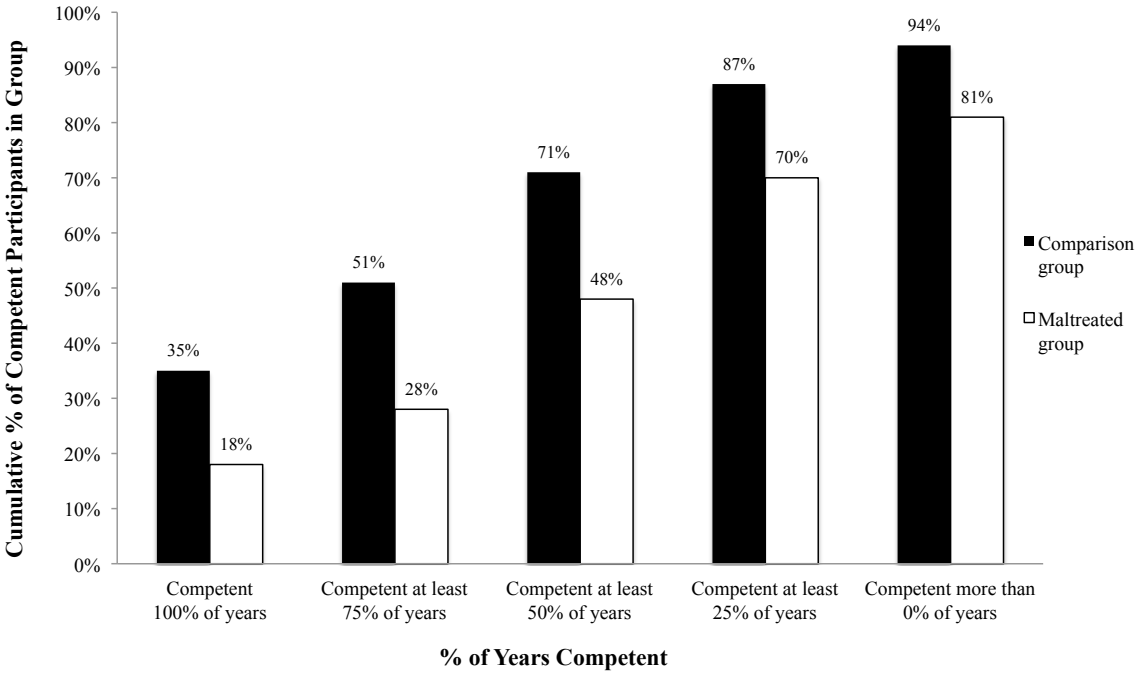

Figure 2: Cumulative Percentage of Years Competent Across Groups

3.1.2 Domains of Competent Functioning

The total number of domains within which participants were consistently competent across all years of data also was calculated for the maltreated group, the comparison group, and the overall sample. The results of these analyses are presented in Table 4. Participants were not penalized for missing data with respect to standardized testing within the academic performance domain, with the result that the number of participants who were consistently competent across

all years of data (n = 188,777) is slightly higher than those who demonstrated consistent competence within all three domains of functioning across all years of available data (n = 184,807). However, the pattern of results is virtually identical, with 17% of the maltreated sample and 35% of the nonmaltreated sample demonstrating competence in all three domains across all years. Similarly, 16% of the maltreated sample and 7% of the nonmaltreated sample were not consistently competent across years in any of the three domains, while a very small percentage of both groups (0.5% of the maltreated group and 0.03% of the nonmaltreated group) did not demonstrate competence in any domain in any year (n = 860). The difference between groups in the number of domains in which participants were consistently competent (not controlling for any covariates) was significant (z = -163.316, $p < .0001$).

Table 4: Number of Domains in Which Participants Were Consistently Competent Across All Years

Group	Never competent in any domain	Consistently competent in 0 domains	Consistently competent in 1 domain	Consistently competent in 2 domains	Consistently competent in 3 domains
Overall sample	860	55,973	171,313	196,339	184,807
Percentage	0.1%	9%	28%	32%	30%
Maltreated group	732	23,845	57,396	45,564	26,692
Percentage	0.5%	16%	37%	30%	17%
Comparison group	128	32,128	113,917	150,775	158,115
Percentage	0.03%	7%	25%	33%	35%

Participants' consistent competence within each specific domain across all available years of data also was calculated for the maltreated group and the comparison group and is shown in Table 5. More participants demonstrated consistent competence in the domain of special education (75% of the overall sample) than behavioral functioning (61%) or academic performance (48%). The nonmaltreated group demonstrated higher rates of consistent competence within each domain of functioning, and the differences between groups with respect

to consistent competence within each specific domain (not controlling for any covariates) were significant: academic performance ($\chi^2(1) = 16{,}367.162$, $p < .0001$); special education ($\chi^2(1) = 3{,}281.854$, $p < .0001$); and behavioral functioning ($\chi^2(1) = 19{,}657.631$, $p < .0001$).

Table 6 sets forth the distribution of consistently competent functioning overall and within each domain of functioning for participants by grade category. In the early education years, covering preschool through 2nd grade, participants demonstrated relatively high rates of competence, with competence rates of 73% to 94% across years. Participants' competence rates were relatively lower in later years of elementary education and junior high, ranging from 41% to 53% for participants in 3rd through 5th grades and ranging from 47% to 53% for most participants in 6th through 8th grades. A small cohort of participants also had junior high data in 2007, with 78% demonstrating competence in that year. In high school, encompassing 9th grade through 11th grade in the sample, participants' rates of competence increased overall, ranging from 60% to 62% for most of the sample. A small cohort of participants also had high school data in 2010, with 67% demonstrating competence in that year.

Table 5: Competence Within and Across Specific Domains of Functioning

Specific domain by group	2006	2007	2008	2009	2010	2011	2012	All years
Academic performance - Total	260,747	315,225	249,210	308,065	426,436	407,062	393,635	291,709
% competent overall	84%	81%	65%	72%	76%	73%	71%	48%
Maltreated group	50,223	59,002	44,571	55,409	84,510	83,234	82,849	51,622
% competent	78%	71%	52%	57%	64%	61%	58%	34%
Comparison group	210,524	256,223	204,639	252,656	341,926	323,828	310,786	240,087
% competent	86%	84%	68%	76%	79%	77%	75%	53%
Special education - Total	306,227	371,791	410,590	478,137	459,956	453,839	453,516	456,136
% competent overall	96%	93%	84%	84%	81%	81%	81%	75%
Maltreated group	63,206	76,840	83,498	97,907	99,229	102,605	106,930	106,668
% competent	95%	90%	78%	77%	75%	75%	75%	69%
Comparison group	243,021	294,951	327,092	380,230	360,727	351,234	346,586	349,468
% competent	96%	94%	86%	85%	84%	83%	83%	77%
Behavioral functioning - Total	312,626	387,075	457,907	492,292	443,319	453,841	436,006	370,567
% competent overall	98%	97%	94%	86%	79%	81%	78%	61%
Maltreated group	63,419	79,561	94,797	98,413	90,293	93,725	91,519	70,310
% competent	95%	93%	89%	77%	68%	68%	64%	46%
Comparison group	249,207	307,514	363,110	393,879	353,026	360,116	344,487	300,257
% competent	98%	98%	96%	88%	82%	85%	83%	66%

Table 6: Consistent Competence Within Domains and Years by Grade Category

Competence by grade category	2006	2007	2008	2009	2010	2011	2012
Total early elementary (Pre-K - 2nd Grade)	206,540	218,883	217,492	103,258	159,558	83,831	9,042
% competent overall	94%	94%	82%	83%	73%	78%	84%
% competent - academic performance	96%	96%	92%	98%	99%	100%	100%
% competent - special education	100%	100%	86%	88%	86%	85%	93%
% competent - behavioral functioning	98%	98%	99%	96%	85%	91%	84%
Total elementary (3rd Grade - 5th Grade)	103,441	173,572	208,635	222,346	228,823	229,343	228,627
% competent overall	56%	56%	41%	47%	47%	52%	53%
% competent - academic performance	62%	63%	47%	62%	65%	66%	67%
% competent - special education	87%	84%	83%	79%	77%	77%	77%
% competent - behavioral functioning	97%	95%	91%	83%	82%	88%	88%
Total junior high (6th Grade - 8th Grade)	-	67	39,396	105,000	175,033	207,431	216,578
% competent overall	-	78%	53%	50%	47%	48%	49%
% competent - academic performance	-	85%	63%	69%	69%	67%	67%
% competent - special education	-	97%	91%	85%	84%	83%	82%
% competent - behavioral functioning	-	93%	82%	73%	68%	71%	72%
Total high school (9th Grade - 11th Grade)	-	-	-	-	72	38,616	102,206
% competent overall	-	-	-	-	67%	62%	60%
% competent - academic performance	-	-	-	-	93%	83%	85%
% competent - special education	-	-	-	-	96%	91%	88%
% competent - behavioral functioning	-	-	-	-	72%	73%	71%

3.1.3 Multilevel Analyses of the Effect of Maltreatment on Consistent Competence

Multilevel logistic regression analyses were used to examine differences in the prevalence of consistent competence between the maltreated group and nonmaltreated group. These analyses were designed to account for possible non-independence of observations resulting from participants clustered at the county level, correcting for the resulting correlated error. In addition, these analyses controlled for the effects of several demographic variables and other covariates, including gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data.

These multilevel analyses were conducted through a set of three fitted models. Model 1 was an unconditional, intercept-only model with one variable at Level 2, participants' school county, predicting the Level 1 intercept of the binary response variable, consistent competence, as a random effect of the Level 2 grouping variable, with no other predictors at Level 1 or Level 2. The variance term for the random effect of county on competence was used to calculate the intra-

class correlation (ICC), a measure of the relatedness of clustered data that compares the amount of variance between clusters to the amount of variance within clusters. For binary outcomes, the within-cluster component of variance is fixed at $\pi^2 / 3$, and the ICC is calculated as: between-cluster (Level 2) variance / (Level 2 variance + $\pi^2 / 3$), or .165 / (.165 + 3.29) = .048 (Hox, 2010; Snijders & Bosker, 1999). The calculated ICC indicated that approximately 5% of the total variance in the odds of a participant being competent could be attributed to differences between counties. Given the large average sample size per cluster ($n \cong 6{,}000$ per cluster), even such a modest amount of county-level variance could introduce error into parameter estimates. Thus, the multilevel structure was deemed appropriate to conduct the planned logistic regression models.

Building upon Model 1, Model 2 introduced a single Level 1 predictor into the model to examine whether being maltreated was associated with the odds of being consistently competent. Participants in the comparison group were found to have significantly greater odds of being competent than participants in the maltreated group, $t(101) = 123.240$, $p < .0001$. The estimated odds ratio for maltreatment status was 2.537, 95% CI [2.499, 2.575], indicating that nonmaltreated participants had 2.5 times the odds of being competent compared to participants who had been maltreated. Fit statistics for Model 2 indicated improved fit over the unconditional model.

Model 3 included several Level 1 demographic variables and covariates to control for the effects of gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data on consistent competence. Maltreatment status continued to predict competence after controlling for the effects of the covariates. The estimated odds ratio for maltreatment status adjusted for covariates was 1.926, 95% CI [1.891, 1.962], indicating that nonmaltreated participants had almost twice the

odds of being competent compared to participants who had been maltreated. Model fit statistics indicated improved fit over Model 2 with maltreatment status as the only Level 1 variable.

As shown in Table 8, all Model 3 covariates were significantly related to the response variable of consistent competence. Female participants had twice the odds of being competent as males. Similarly, White participants had almost 3 times the odds of being competent as Hispanic participants and more than twice the odds of being competent as Black participants. Participants from high SES families had 4 times the odds of being competent than participants from low SES families, and middle SES participants had almost twice the odds of being competent as low SES participants. The odds of being competent were greater for participants who were born full-term with normal birth weight than those born prematurely or with low birth weight, and participants whose mothers were married at the time of their birth also had greater odds of being competent than those born to mothers who were not married. Having fewer years of education data or having a later birth year (being younger) was related to greater odds of being competent. Mother's age at birth also was modestly yet significantly related to competence, with participants with older mothers having slightly greater odds of being competent.

Table 7 presents the fixed effects, random parameters, and model fit statistics for each of the three fitted models. Table 8 describes the odds ratios and 95% confidence intervals for each fixed effect in these models.

Table 7: Fixed Effects and Covariance Estimates for Multilevel Models Predicting Consistent Competence for Children Nested Within School County

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-.821*** (.041)	-1.534*** (.042)			-2.360*** (.039)		
Maltreatment status			(1, 101)	15,187.700***		(1, 101)	5,000.050***
Comparison		.931*** (.008)			.656*** (.009)		
Maltreated		Reference			Reference		
Gender						(1, 101)	11,881.000***
Female					.711*** (.007)		
Male					Reference		
Race/ethnicity						(3, 300)	3,372.420***
Black					-.868*** (.010)		
Hispanic					-1.044*** (.016)		
Other					-.242*** (.012)		
White					Reference		
SES						(2, 202)	9,890.460***
High SES					1.393*** (.010)		
Middle SES					.638*** (.009)		
Low SES					Reference		
Fetal well-being						(1, 100)	222.030***
Full-term and normal BW					.150*** (.010)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	388.180***
Married					.175*** (.009)		
Not married					Reference		
Birth year (GMC)					.178*** (.002)	(1, 608351)	11,448.700***
Mother's age at birth (GMC)					.006*** (.001)	(1, 608351)	95.570***
Years of education data (GMC)					-.281*** (.002)	(1, 608351)	13,908.700***
Level 2		Random parameters					
County-level variance	.165 (.025)	.171 (.025)			.132 (.019)		
-2LL (random intercept)	753,592.000	735,661.000			615,513.000		
Chi-square (random intercept)	32,364.400***	31,279.900***			31,237.700***		
-2LL (random effects)	720,653.000	703,806.000			583,748.600		
Pearson chi-square	608,027.000	614,680.000			597,058.900		
Pearson chi-square / DF	1.000	1.010			0.980		
Model fit statistics							
-2LL	721227.000	704381.000			584275.800		
AIC	721231.000	704387.000			584303.800		
BIC	721237.000	704395.000			584340.600		

Note. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor, maltreatment status. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.

***All parameter estimates were statistically significant at $p < .001$.

Table 8: Odds Ratios for Multilevel Models Predicting Consistent Competence for Children Nested Within School County

Variable	Model 2		Model 3	
	OR	95% CI	*OR / IOR*	95% CI
Maltreatment status				
Comparison vs. maltreated	2.537	[2.499, 2.575]	1.926	[1.891, 1.962]
Gender				
Female vs. male			2.036	[2.010, 2.063]
Race/ethnicity				
White vs. Black			2.381 *(IOR)*	[2.336, 2.427]
White vs. Hispanic			2.841 *(IOR)*	[2.755, 2.933]
White vs. Other			1.274 *(IOR)*	[1.244, 1.304]
SES				
High SES vs. low SES			4.028	[3.949, 4.108]
Middle SES vs. low SES			1.893	[1.859, 1.928]
Fetal well-being				
Full-term and normal BW vs. premature and low BW			1.162	[1.139, 1.185]
Mother's marital status				
Married vs. not married			1.191	[1.170, 1.212]
Birth year (GMC)			1.195	[1.191, 1.198]
Mother's age at birth (GMC)			1.006	[1.005, 1.008]
Years of education data (GMC)			1.325 *(IOR)*	[1.318, 1.332]

Note. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor, maltreatment status. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

3.1.4 Multilevel Analyses of the Effect of Maltreatment on Patterns of Competence

In addition to the primary multilevel analyses investigating the effect of maltreatment status on participants' consistently competent functioning across time, supplemental multilevel analyses examined the effects of maltreatment on patterns of competence among participants. Multilevel logistic regression models were used to analyze consistent competence across time within specific domains of functioning, including academic performance, special education, and behavioral domains. Multilevel regression analyses were used to test the effects of maltreatment on the percentage of years in which participants demonstrated competence across all domains.

After controlling for the effects of covariates, including gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data, maltreatment status predicted competence across time in the domain of academic functioning, $F(1, 101) = 2{,}735.380$, $p < .0001$. Nonmaltreated participants had 1.5 times the odds of functioning competently in academic performance than maltreated participants ($OR = 1.503$, 95% CI [1.480, 1.526]). Results were similar with respect to competence over time within the special education domain ($F(1, 101) = 2{,}735.380$, $p < .0001$; $OR = 1.460$, 95% CI [1.437, 1.483]). The effect of maltreatment was slightly larger in the behavioral functioning domain, $F(1, 101) = 2{,}303.820$, $p < .0001$, with nonmaltreated participants having more than twice the odds of functioning competently over time with respect to behavior as maltreated participants ($OR = 2.118$, 95% CI [2.084, 2.151]). In multilevel regression analyses including maltreatment status as a Level 1 predictor and the full set of covariates, maltreatment status also predicted the percentage of years that participants were consistently competent, $F(1, 101) = 13{,}557.400$, $p < .0001$.

3.2 Multilevel Analyses of the Effect of the Introduction of Multiple Response System Within Counties

Multilevel logistic regression analyses were used to examine the effect of the introduction of the Multiple Response System within each maltreated participant's county of residence. These analyses used only the maltreated group and examined whether participants who were investigated or assessed for maltreatment after the date of the introduction of the Multiple Response System in their county were more likely to be resilient following experiences of maltreatment.

Though most of the maltreated participants were reported for maltreatment and educated within the same county, approximately 13.5% of the maltreated group appeared to have moved at some point after the date of their CPS investigation and attended school in a different county for all or most of their education years used in the current study. Without specific information about the dates on which these participants' families may have moved, it is difficult to determine how long these participants may have benefitted from child welfare reform in the county in which they were reported for maltreatment. As a result, three separate sets of multilevel analyses were performed on the following samples of participants: (1) the full maltreated group (n = 153,497), (2) the subsample of maltreated participants who appeared to have remained in the same county for school after being reported for maltreatment (n = 132,828), and (3) the subsample of maltreated participants who appeared to have moved to a different county for all or most of their available school years (n = 20,667). These multilevel analyses controlled for possible non-independence of observations among participants within the same county as well as the effects of several demographic variables and other covariates.

The relative percentage of maltreated participants who were reported for maltreatment prior to adoption of the Multiple Response System in their county was 53% in the full maltreated sample, 52% in the nonmover subsample, and 62% in the mover subsample. Prior to conducting

multilevel analyses, the relative percentages of resilient and non-resilient participants who were reported for maltreatment prior to or after the introduction of the Multiple Response System in their county were calculated. A significantly higher percentage of resilient participants (63%) were reported for maltreatment after the policy change was implemented in their county, compared to 44% of non-resilient participants (not controlling for any covariates) ($\chi^2(1) = 3{,}350.651$, $p < .0001$).

3.2.1 Multilevel Analyses of the Effect of Introduction of Multiple Response System on Consistent Competence

Table 9, Table 10, and Table 11 present the fixed effects, random parameters, and model fit statistics for each of three fitted models investigating the effect of the introduction of the Multiple Response System on consistent competence for participants in the full maltreated sample, the subsample of nonmovers, and the subsample of movers, respectively. Table 12 describes the odds ratios and 95% confidence intervals for each fixed effect in these models. The models will be discussed below with respect to each sample of maltreated participants used in the analyses.

Model 1 was an unconditional, intercept-only model with one variable at Level 2, participants' school county, predicting the Level 1 intercept of the binary response variable, consistent competence following experiences of maltreatment, as a random effect of the Level 2 grouping variable, with no other predictors at Level 1 or Level 2. The variance term for the random effect of county on consistent competence for the full maltreated sample was .130, indicating that approximately 4% of the total variance in the odds of a participant being competent could be attributed to differences between counties. The ICC was identical for the subsample of nonmovers within the maltreated sample. For the subsample of maltreated participants who moved counties sometime after their maltreatment report was investigated, the

variance term was smaller, .079, with approximately 2% of the total variance in the odds of a participant being competent being attributable to differences between counties.

Model 2 introduced a single Level 1 predictor into the model to examine whether being reported for maltreatment in a county that had adopted the Multiple Response System increased the likelihood of competent outcomes. Participants who were reported to CPS after the Multiple Response System was adopted in their county were found to have significantly greater odds of being competent than participants reported prior to the introduction of the Multiple Response System, $t(101) = 61.290$, $p < .0001$. The estimated odds ratio associated with the introduction of the Multiple Response System was 2.402, 95% CI [2.335, 2.471], indicating that maltreated participants who were investigated for maltreatment in a county that had adopted the child welfare policy change had almost 2.5 times the odds of being competent compared to maltreated participants who lived in a county that had not yet adopted the policy at the time their maltreatment report was investigated. Fit statistics for Model 2 indicated improved fit over the unconditional model. As shown in Table 12, the pattern of results was similar across nonmover and mover subsamples (*OR* = 2.458, 95% CI [2.382, 2.535], for nonmovers; *OR* = 2.241, 95% CI [2.083, 2.411], for movers).

Model 3 included several Level 1 demographic variables and covariates to control for the effects on consistent competence of several factors, including gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data. Being reported for maltreatment in a county that had adopted the Multiple Response System continued to predict subsequent consistent competence after controlling for the effects of the covariates. The estimated odds ratio adjusted for covariates for the full maltreated sample was 2.265, 95% CI [2.199, 2.333], indicating that maltreated participants investigated or assessed by CPS in Multiple Response System counties

had more than twice the odds of being competent compared to participants reported to CPS prior to adoption of the policy in their county. Model fit statistics across samples indicated improved fit over Model 2 with the Multiple Response System predictor as the only Level 1 variable. As shown in Table 9, Table 10, Table 11, and Table 12, the model parameter estimates and odds ratios were similar across subsamples of movers and nonmovers.

All Model 3 covariates except for mother's marital status were significantly related to consistent competence. As shown in Table 12, female participants in the maltreated sample had twice the odds of being competent as males, after adjusting for the other variables in the model. Similarly, White maltreated participants had more than twice the odds of being competent as Hispanic maltreated participants and Black maltreated participants. Maltreated participants from high SES families had more than three times the odds of being competent as maltreated participants from low SES families, and middle SES participants in the maltreated sample had twice the odds of being competent as low SES participants in the maltreatment sample. The odds of being competent were greater for maltreated participants born full-term at a normal birth weight than maltreated participants who were born prematurely or had a low birth weight. Maltreated participants who were younger or had fewer years of education data had greater odds of being competent, as evidenced by the significant parameter estimate associated with the cross-product of these two indicators.

Mother's age at the time of birth was found to have a modest yet significant negative association with consistent competence, with participants born to younger mothers having slightly greater odds of being competent. The effect of mother's age in the full maltreated sample was quite small (inverse odds ratio *(IOR)* = 1.005, 95% CI [1.002, 1.007]) and does not seem to indicate a meaningful difference in participants' likelihood of being competent based on mother's age once other variables in the model have been taken into account. Examination of the

unadjusted relation between mother's age at birth and consistent competence demonstrates a small and significant positive association (r = .062 for maltreated sample, r = .176 for full sample). There was no significant difference in consistent competence among maltreated participants based on mother's marital status at the time of birth after controlling for the effects of the Level 1 predictor and the other covariates.

Supplemental analyses investigated the effects of one component of the Multiple Response System, the implementation of family assessments within specific counties. Family assessments were an important part of the Multiple Response System. Multilevel logistic regression models compared the date on which a participant was investigated or assessed for maltreatment to the date that participant's county began reporting family assessments to the state to test whether formal implementation of family assessments within the county investigating or assessing participant's maltreatment report predicted their consistent competence. The results of multilevel models that compared participants reported for maltreatment after the initial use of family assessments within their county to maltreated participants reported to CPS before family assessments began were nearly identical to those in the foregoing analyses involving the introduction of the Multiple Response System as a whole and are not reported separately.

Table 9: Fixed Effects and Covariance Estimates for Multilevel Models Predicting Competence for Maltreated Children Nested Within School County (Multiple Response System – Full Sample)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-1.471*** (.038)	-1.924*** (.040)			-2.522*** (.041)		
Multiple Response System			(1, 101)	3,756.840***		(1, 101)	2,987.250***
After policy introduced		.876*** (.014)			.818*** (.015)		
Before policy introduced		Reference			Reference		
Gender						(1, 101)	2,428.430***
Female					.715*** (.015)		
Male					Reference		
Race/ethnicity						(3, 290)	579.790***
Black					-.812*** (.020)		
Hispanic					-.658*** (.039)		
Other					-.164*** (.024)		
White					Reference		
SES						(2, 200)	1,634.050***
High SES					1.24*** (.024)		
Middle SES					.716*** (.017)		
Low SES					Reference		
Fetal well-being						(1, 100)	15.090***
Full-term and normal BW					.080*** (.021)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	1.620 *ns*
Married					.021*** (.016)		*p* = .206
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.058*** (.002)	(1, 153479)	1,308.640***
Mother's age at birth (GMC)					-.005*** (.001)	(1, 153479)	11.330***
Level 2		Random parameters					
County-level variance	.130 (.020)	.136 (.021)			.092 (.015)		
-2LL (random intercept)	143,952.000	140,587.000			128,873.000		
Chi-square (random intercept)	3,041.700***	3,756.840***			2,379.590***		
-2LL (random effects)	140,528.000	136,629.000			126,155.200		
Pearson chi-square	153,017.900	151,937.700			154,032.200		
Pearson chi-square / DF	1.000	0.990			1.000		
Model fit statistics							
-2LL	140,910.100	137,013.000			126,493.400		
AIC	140,914.100	137,019.000			126,519.400		
BIC	140,919.400	137,026.900			126,553.600		

*Note. S*ample used in analyses was full sample of maltreated participants. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects whether a participant was investigated for maltreatment before or after introduction of the Multiple Response System in the county investigating their maltreatment report. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.

***All parameter estimates were statistically significant at $p < .001$ except as noted. *ns* = not significant.

Table 10: Fixed Effects and Covariance Estimates for Multilevel Models Predicting Competence for Maltreated Children Nested Within School County (Multiple Response System - Nonmovers)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-1.472*** (.040)	-1.950*** (.042)			-2.525*** (.044)		
Multiple Response System			(1, 198)	3,300.030***		(1, 98)	2,588.130***
After policy introduced		.899*** (.016)			.834*** (.016)		
Before policy introduced		Reference			Reference		
Gender						(1, 98)	2,013.100***
Female					.702*** (.016)		
Male					Reference		
Race/ethnicity						(3, 279)	529.060***
Black					-.838*** (.022)		
Hispanic					-.661*** (.041)		
Other					-.156*** (.026)		
White					Reference		
SES						(2, 193)	1,450.650***
High SES					1.252*** (.026)		
Middle SES					.731*** (.018)		
Low SES					Reference		
Fetal well-being						(1, 98)	11.940***
Full-term and normal BW					.077*** (.022)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 98)	0.210 *ns*
Married					.008*** (.018)		*p* = .649
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.058*** (.002)	(1, 132814)	1,308.640***
Mother's age at birth (GMC)					-.004*** (.002)	(1, 132814)	11.330***
Level 2		Random parameters					
County-level variance	.140 (.023)	.144 (.023)			.094 (.016)		
-2LL (random intercept)	124,283.000	121,372.000			110,892.000		
Chi-square (random intercept)	2,874.030***	3,405.540***			2,234.130***		
-2LL (random effects)	121,043.400	117,600.700			108,340.000		
Pearson chi-square	132,363.000	131,206.100			132,994.200		
Pearson chi-square / DF	1.000	0.990			1.000		
Model fit statistics							
-2LL	121,409.300	117,966.700			108,658.100		
AIC	121,413.300	117,972.700			108,684.100		
BIC	121,418.500	117,980.500			108,717.800		

Note. Sample used in analyses was subsample of maltreated participants investigated for maltreatment in the same county in which they attended school for most or all years of available education data. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects whether a participant was investigated for maltreatment before or after the introduction of the Multiple Response System in their county. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.
***All parameter estimates were statistically significant at $p < .001$ except as noted. *ns* = not significant.

Table 11: Fixed Effects and Covariance Estimates for Multilevel Models Predicting Competence for Maltreated Children Nested Within School County (Multiple Response System - Movers)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-1.480*** (.036)	-1.840*** (.041)			-2.563*** (.074)		
Multiple Response System			(1, 101)	478.810***		(1, 101)	396.610***
After policy introduced		.807*** (.037)			.763*** (.038)		
Before policy introduced		Reference			Reference		
Gender						(1, 101)	419.990***
Female					.797*** (.039)		
Male					Reference		
Race/ethnicity						(3, 262)	54.710***
Black					-.650*** (.052)		
Hispanic					-.601*** (.139)		
Other					-.207*** (.062)		
White					Reference		
SES						(2, 197)	174.270***
High SES					1.117*** (.070)		
Middle SES					.622*** (.044)		
Low SES					Reference		
Fetal well-being						(1, 99)	3.010***
Full-term and normal BW					.093*** (.054)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	3.350 *ns*
Married					.077*** (.042)		*p* = .070
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.058*** (.005)	(1, 20552)	163.800***
Mother's age at birth (GMC)					-.009*** (.004)	(1, 20552)	5.270*
Level 2		Random parameters					
County-level variance	.079 (.019)	.082 (.020)			.062 (.017)		
-2LL (random intercept)	19,665.000	19,179.000			17,918.000		
Chi-square (random intercept)	93.480***	88.690***			60.960***		
-2LL (random effects)	19,403.390	18,921.100			17,715.360		
Pearson chi-square	20,396.540	20,359.920			20,746.390		
Pearson chi-square / DF	0.990	0.990			1.000		
Model fit statistics							
-2LL	19,571.700	19,090.300			17,856.850		
AIC	19,575.700	19,096.300			17,882.850		
BIC	19,580.950	19,104.180			17,916.970		

Note. Sample used was subsample of maltreated participants reported for maltreatment in a different county than the county in which they attended school for most or all years of available education data. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects whether a participant was investigated for maltreatment before or after the introduction of the Multiple Response System in their county. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion. ***All parameter estimates were statistically significant at $p < .001$ except as noted. *$p < .05$. *ns* = not significant.

Table 12: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Competence for Maltreated Children Nested Within School County (Multiple Response System)

Variable	Full maltreated sample		Nonmovers		Movers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
Multiple Response System						
After vs. before policy introduced	2.402	[2.335, 2.471]	2.458	[2.382, 2.535]	2.241	[2.083, 2.411]
			Model 3			
Multiple Response System						
After vs. before policy introduced	2.265	[2.199, 2.333]	2.303	[2.229, 2.379]	2.145	[1.988, 2.314]
Gender						
Female vs. male	2.043	[1.985, 2.103]	2.017	[1.955, 2.080]	2.218	[2.054, 2.396]
Race/ethnicity						
White vs. Black	2.252	[2.165, 2.342]	2.309 *(IOR)*	[2.212, 2.415]	1.916	[1.727, 2.123]
White vs. Hispanic	1.931	[1.789, 2.083]	1.938 *(IOR)*	[1.789, 2.096]	1.825	[1.387, 2.398]
White vs. Other	1.178	[1.124, 1.236]	1.170 *(IOR)*	[1.110, 1.232]	1.230	[1.088, 1.389]
SES						
High SES vs. low SES	3.444	[3.283, 3.612]	3.496	[3.322, 3.679]	3.055	[2.659, 3.509]
Middle SES vs. low SES	2.046	[1.979, 2.115]	2.076	[2.033, 2.153]	1.862	[1.706, 2.033]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.083	[1.040, 1.128]	1.080	[1.033, 1.129]	1.098	[.987, 1.221]
Mother's marital status						
Married vs. not married	1.021	[.989, 1.054]	1.008	[.973, 1.044]	1.080	[.994, 1.173]
Birth year (GMC, RC) X years of education data (GMC)	1.059	[1.056, 1.063]	1.059	[1.056, 1.063]	1.060	[1.051, 1.069]
Mother's age at birth (GMC)	1.005 *(IOR)*	[1.002, 1.007]	1.004 *(IOR)*	[1.001, 1.007]	1.009 *(IOR)*	[1.001, 1.017]

Note. Samples used in analyses were: full sample of maltreated participants, subsample of maltreated group who were reported for maltreatment in the same county in which they attended school for most or all years of available education data, and subsample of maltreated group who were reported for maltreatment in a different county than the county in which they attended school for most or all years of available education data. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects whether a participant was investigated for maltreatment before or after the introduction of the Multiple Response System in the county investigating the maltreatment report. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being resilient as comparison group).

3.2.2 Multilevel Analyses of the Effect of Introduction of Multiple Response System on Patterns of Competence

In addition to the primary multilevel analyses investigating the effect of the introduction of the Multiple Response System on maltreated participants' consistently competent functioning across time, supplemental multilevel analyses examined the effects of the introduction of the Multiple Response System on patterns of competence among maltreated participants. Multilevel logistic regression models were used to analyze effects on consistent competence across time within specific domains of functioning, including academic performance, special education, and behavioral functioning domains. Multilevel regression analyses were used to test effects on the percentage of years in which maltreated participants demonstrated competence across all domains.

After controlling for the effects of covariates, including gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data, the introduction of the Multiple Response System within participants' county prior to the date their maltreatment report was investigated predicted competence across time in the domain of academic functioning, $F(1, 101) = 4{,}998.970$, $p < .0001$. Maltreated participants who were investigated for maltreatment after the Multiple Response System was adopted in their county had more than twice the odds of functioning competently in academic performance than maltreated participants who were investigated for maltreatment in a county that had not yet adopted the Multiple Response System ($OR = 2.335$, 95% CI [2.280, 2.391]). Results were similar but demonstrated a slightly smaller effect with respect to the behavioral functioning domain, $F(3, 101) = 3{,}077.770$, $p < .0001$, $OR = 1.924$, 95% CI [1.880, 1.970]). The effect of the introduction of the Multiple Response System was smaller but still significant with respect to competence over time within the special education domain ($F(1, 101) = 266.740$, $p < .0001$; $OR = 1.218$, 95% CI [1.190, 1.248]). In multilevel regression analyses including the introduction of the Multiple Response System as a Level 1 predictor as

well as the covariates, being investigated for maltreatment in a county that had already adopted the Multiple Response System predicted a higher percentage of years of consistent competence, $F(1, 101) = 74.900$, $p < .0001$.

3.3 Multilevel Analyses of the Effect of Government Expenditures on Early Childhood Programs

Multilevel logistic regression analyses were used to examine whether levels of expenditures on Smart Start and More at Four predicted consistent competence among maltreated participants. These analyses also examined whether the effects of government expenditures on these early childhood programs operated similarly across the maltreated and nonmaltreated groups. These multilevel analyses were designed to account for possible non-independence of observations resulting from participants clustered at the county level, correcting for the resulting correlated error. In addition, these analyses controlled for the effects of several demographic variables and other covariates.

The following sets of analyses were conducted with respect to the effects of expenditures on Smart Start and More at Four on participants' consistent competence across domains and years. First, the overall effect of expenditures for these programs was calculated for the full sample of participants. Second, the interaction of maltreatment status and level of government expenditures on these early childhood programs was examined to determine whether there were any differential impacts on the maltreated group. Third, separate analyses were conducted with respect to participants in the maltreated group and participants in the comparison group to compare the specific pattern of effects across groups.

Additional analyses also were conducted to determine whether findings were robust to assumptions regarding participants' county of residence between birth and age 5. Though most participants were born and educated within the same county, 26% of the maltreated group and

24% of the comparison group appeared to have moved at some point after their birth and attended school in a different county for all or most education years used in the current study. Without specific information about the dates on which these participants' families may have moved, it is difficult to determine how long these participants may have benefitted from government expenditures for early childhood programs in the county in which they were born.

As a result of these apparent moves, separate multilevel analyses were performed on three subsamples of participants within the overall sample and within each group: (1) the full sample or group, including those who moved (referred to as "movers"), assuming all participants remained in their birth county during the ages of 0-5 and thus received the benefit of Smart Start and More at Four entirely in their birth county, (2) the full sample or group, assuming all movers moved immediately after birth to the county in which they attended school for most or all of their education years and thus received the benefit of Smart Start and More at Four entirely in their school county, and (3) the subsample of participants who remained in the same county for school after their birth (referred to as "nonmovers"). For the full sample, the total sample including movers and nonmovers was n = 608,432, and the total nonmover sample was n = 460,036. For the maltreated group, the total sample including movers and nonmovers was n = 153,497, and the total nonmover sample was n = 113,521. For the comparison group, the total sample including movers and nonmovers was n = 454,935, and the total nonmover sample was n = 346,515.

Total per capita expenditures on Smart Start and More at Four within counties were categorized as small (less than $1,500, approximately 24% of each group), moderate (between $1,500 and $2,500, approximately 52% of the maltreated group and 56% of the comparison group), and large (greater than $2,500, approximately 24% of the maltreated group and 20% of the comparison group). These relative percentages were consistent across the three different subsamples of movers and nonmovers within each group used in these analyses.

Prior to conducting multilevel analyses, the relative percentages of competent and non-competent participants at each level of county expenditures on Smart Start and More at Four were calculated. Among participants living in counties with small expenditures on Smart Start and More at Four, 26% were competent, whereas among participants living in counties with large expenditures on Smart Start and More at Four, 34% were competent. Thirty-two percent (32%) of participants living in counties with moderate expenditures on Smart Start and More at Four were competent. The overall effect of expenditures on Smart Start and More at Four on competence in the full sample (without any covariates) was significant ($\chi^2(2) = 2{,}073.693$, $p < .0001$).

In the maltreated group, 14% of participants living in counties with small expenditures on Smart Start and More at Four were competent, whereas 22% of participants living in counties with large expenditures on Smart Start and More at Four were competent. Eighteen percent (18%) of participants living in counties with moderate expenditures on Smart Start and More at Four were competent. Relative levels of government expenditures predicted competence among maltreated participants ($\chi^2(2) = 760.689$, $p < .0001$).

In the comparison group, 31% of participants living in counties with small expenditures on Smart Start and More at Four were competent, whereas 39% of participants living in counties with large expenditures on these early childhood programs were competent. Thirty-six percent (36%) of participants living in counties with moderate expenditures on Smart Start and More at Four were competent. Relative levels of government expenditures predicted competence ($\chi^2(2) = 1{,}637.444$, $p < .0001$).

The effect of county expenditures for Smart Start and More at Four on consistent competence was examined for the full sample as well as for each of the maltreated group and comparison group separately to determine whether there was any differential impact on

maltreated participants. The models will be discussed below with respect to each sample of participants used in the analyses.

3.3.1 Full Sample Including Movers in Birth County

The overall effect of government expenditures on Smart Start and More at Four was examined using multilevel logistic regression. The first set of analyses included the full sample of maltreated and nonmaltreated participants assuming movers remained in their birth county during the ages of 0-5 and thus received the benefit of expenditures on Smart Start and More at Four in their birth county. Model 1 was an unconditional, intercept-only model with one variable at Level 2, participants' school county, predicting the Level 1 intercept of the binary response variable, consistent competence, as a random effect of the Level 2 grouping variable, with no other predictors at Level 1 or Level 2. The variance term for the random effect of county on consistent competence in Model 1 for the full sample assuming that movers remained in their birth county was .170, indicating that approximately 5% of the total variance in the odds of a participant being consistently competent could be attributed to differences between counties.

Model 2 included a Level 1 predictor to examine whether the amount of government expenditures on early childhood programs was associated with the odds of being consistently competent. In Model 2, participants whose birth counties were allocated higher amounts of funding for Smart Start and More at Four during the years in which participants were age 0 to 5 were found to have significantly greater odds of being competent than participants whose birth counties received lower amounts of funding, $F(2, 202) = 3{,}350.690$, $p < .0001$. The estimated odds ratio for large compared to small expenditures on early childhood programs was 2.238, 95% CI [2.190, 2.287], indicating that participants whose birth counties received relatively large allocations for expenditures on these programs had more than twice the odds of being consistently competent compared to participants whose birth counties received relatively small allocations.

The estimated odds ratio for moderate compared to small allocations for early childhood programs was 1.792, 95% CI [1.764, 1.821]. Fit statistics for Model 2 indicated improved fit over the unconditional model.

Model 3 included maltreatment status as an additional Level 1 predictor of consistent competence as well as the interaction of level of county expenditures by maltreatment status to investigate whether program effects differed across maltreated and comparison groups. Model 3 also included several Level 1 demographic variables and covariates to control for the effects on consistent competence of gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year (age), and years of education data.

County expenditures on Smart Start and More at Four continued to predict competence after controlling for the effects of the other Level 1 predictors and covariates included in Model 3. The estimated odds ratio for large compared to small expenditures on Smart Start and More at Four adjusted for covariates was 2.593, 95% CI [2.522, 2.665], indicating that participants born in counties with relatively large expenditures on early childhood programs had over 2.5 times the odds of functioning competently across time compared to participants born in counties with relatively small expenditures on early childhood programs. The estimated odds ratio for moderate compared to small expenditures on Smart Start and More at Four was 1.906, 95% CI [1.865, 1.948], indicating that participants in counties with moderate expenditures on early childhood programs had about twice the odds of being competent as participants in counties with small expenditures on early childhood programs.

In addition, the estimated odds ratio for maltreatment status was 1.468, 95% CI [1.441, 1.494], indicating that participants in the comparison group had 1.5 times the odds of being competent as participants in the maltreated group after controlling for the effects of county

expenditures on early childhood programs and adjusting for covariates. The interaction of county expenditures and maltreatment status was significant but quite small, $F(2, 202) = 3.94$, $p = .021$. Model fit statistics indicated improved fit over Model 2 with county expenditures as the only Level 1 variable.

The statistical significance of the interaction term indicated that there were differences between the maltreated group and comparison group with respect to the effect of county expenditures on participants' competent functioning over time. However, given the small size of the effect, it is likely that these differences are not substantial. Figure 3 depicts the interaction of Smart Start and More at Four expenditures by maltreatment status on participants' consistent competence without adjusting for any covariates. Subsequent analyses examined the effects of county expenditures on Smart Start and More at Four on consistent competence within the maltreated group and the comparison group separately to investigate whether there were any differential impacts on maltreated participants.

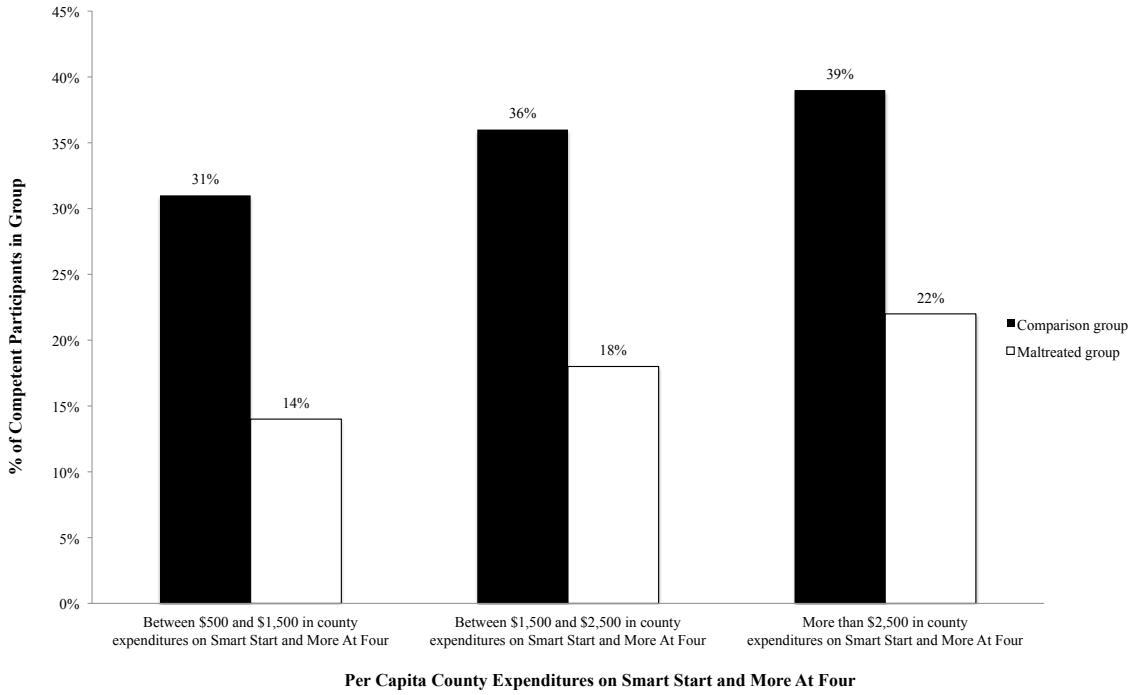

Figure 3: Consistent Competence by Expenditures on Smart Start and More at Four and Maltreatment Status

3.3.2 Full Sample Including Movers in School County

A second set of analyses included the full sample of maltreated and nonmaltreated participants assuming movers lived in their school county after birth and thus received the benefit of expenditures on Smart Start and More at Four in their school county. The Model 1 variance term for the random effect of county on consistent competence for the full sample assuming that movers lived in their school county immediately following birth was the same as in the previous analysis (.170), as the only difference between the two sets of analyses that included movers involved the Level 1 predictor. In Model 2, which added the Level 1 predictor to the unconditional model, participants whose counties received higher amounts of funding for Smart Start and More at Four during the years in which participants were age 0 to 5 were found to have significantly greater odds of being competent than participants whose counties received lower amounts of funding, $F(2, 202) = 4{,}715.860$, $p < .0001$.

The estimated odds ratio for large compared to small expenditures on early childhood programs was 3.323, 95% CI [3.237, 3.412], indicating that participants whose counties received relatively large allocations for expenditures on these programs had more than 3 times the odds of being competent compared to participants whose counties received relatively small allocations. The estimated odds ratio for moderate compared to small expenditures on early childhood programs was 2.058, 95% CI [2.023, 2.094]. Fit statistics for Model 2 indicated improved fit over the unconditional model.

Smart Start and More at Four expenditures continued to predict consistent competence after controlling for the effects of maltreatment status, the interaction of government expenditures and maltreatment status, and the Model 3 covariates. The estimated odds ratio for large compared to small expenditures on Smart Start and More at Four adjusted for covariates and maltreatment status was 4.014, 95% CI [3.887, 4.144], indicating that participants in counties with relatively large expenditures on early childhood programs had 4 times the odds of being competent compared to participants in counties with relatively small expenditures on early childhood programs. The estimated odds ratio for moderate compared to small investments in early childhood programs was 2.334, 95% CI [2.280, 2.389].

The estimated odds ratio for maltreatment status was 1.478, 95% CI [1.441, 1.494], indicating that participants in the comparison group had 1.5 times the odds of being competent as participants in the maltreated group after controlling for the effects of county expenditures on early childhood programs and adjusting for covariates. The interaction of county expenditures and maltreatment status was not significant, $F(2, 196) = 1.39$, $p = .251$, indicating that were no significant differences between the maltreated group and comparison group with respect to the effect of county expenditures on participants' competent functioning over time. Model fit statistics indicated improved fit over Model 2.

3.3.3 Nonmovers Sample

A third set of analyses included only the nonmovers in the full sample. The Model 1 variance term for the random effect of county on consistent competence for the nonmover subsample was .171, indicating that approximately 5% of the total variance in the odds of a participant being competent could be attributed to differences between counties. In Model 2 including government expenditures on early childhood programs as a Level 1 predictor, participants whose birth counties were allocated higher amounts of funding for Smart Start and More at Four during the years in which participants were age 0 to 5 were found to have significantly greater odds of being competent than participants whose birth counties received lower amounts of funding, $F(2, 195) = 3{,}690.320$, $p < .0001$. The estimated odds ratio for large compared to small expenditures on early childhood programs was 3.700, 95% CI [3.583, 3.820], indicating that nonmovers whose counties received relatively large allocations for expenditures on these programs had almost 4 times the odds of being consistently competent compared to nonmovers whose counties received relatively small allocations. The estimated odds ratio for moderate compared to small expenditures on early childhood programs was 2.114, 95% CI [2.071, 2.159]. Fit statistics for Model 2 indicated improved fit over the unconditional model.

In Model 3, county expenditures on Smart Start and More at Four continued to predict competence after controlling for the effects of maltreatment status and the covariates. The estimated odds ratio for large compared to small expenditures on Smart Start and More at Four adjusted for covariates was 4.650, 95% CI [4.473, 4.835], indicating that nonmovers living in counties with relatively large expenditures on early childhood programs had more than 4.5 times the odds of being competent compared to nonmovers living in counties with relatively small expenditures on early childhood programs. The estimated odds ratio for moderate compared to small expenditures on early childhood programs was 2.441, 95% CI [2.373, 2.511]. The estimated

odds ratio for maltreatment status adjusted for covariates was 1.476, 95% CI [1.445, 1.508], indicating that participants in the comparison group had 1.5 times the odds of being competent as participants in the maltreated group after controlling for the effects of county expenditures on early childhood programs and adjusting for covariates. The interaction of county expenditures and maltreatment status was not significant, $F(2, 177) = 1.76$, $p = .175$, indicating that were no significant differences between the maltreated group and comparison group with respect to the effect of county expenditures on participants' competent functioning over time. Model fit statistics indicated improved fit over Model 2 with county expenditures as the only Level 1 variable.

3.3.4 Subgroup Analyses Including Movers in Birth County

Analyses using the full sample assuming that movers lived in their birth county suggested that there might be differences between the maltreated group and comparison group with respect to the effect of county expenditures on participants' competent functioning over time. However, the effect was small, indicating that these differences are not likely to be substantial. Given the focus of the current study on the resilient functioning of maltreated children and in the interest of completeness, the effects of county expenditures on Smart Start and More at Four on consistent competence within the maltreated group and the comparison group were analyzed separately to investigate whether there were any differential impacts on maltreated participants. Table 13, Table 14, and Table 15 present the fixed effects, random parameters, and model fit statistics for each of three fitted models for the full sample of maltreated participants assuming movers remained in their birth county in the relevant years, the full sample of maltreated participants assuming movers moved to their school county immediately after their birth, and the subsample of nonmovers in the maltreated group, respectively. Table 16 describes the odds ratios and 95% confidence intervals for each fixed effect in these models.

The effect of county expenditures for Smart Start and More at Four on consistent competence also was examined for participants in the comparison group to determine whether there was any differential impact on maltreated participants. Table 17, Table 18, and Table 19 present the fixed effects, random parameters, and model fit statistics for each of three fitted models for the full comparison group assuming movers remained in their birth county in the relevant years, the full comparison group assuming movers moved to their school county immediately after their birth, and the subsample of nonmovers in the comparison group, respectively. Table 20 describes the odds ratios and 95% confidence intervals for each fixed effect in these models. The models will be discussed below with respect to each sample of participants used in the analyses.

Model 1 was an unconditional, intercept-only model with one variable at Level 2, participants' school county, predicting the Level 1 intercept of the binary response variable, consistent competence, as a random effect of the Level 2 grouping variable, with no other predictors at Level 1 or Level 2. The variance term for the random effect of county on competence in Model 1 for the full maltreated group assuming that movers remained in their birth county was .130, indicating that approximately 4% of the total variance in the odds of a participant being competent could be attributed to differences between counties. The variance term was similar but somewhat larger for the comparison group assuming movers remained in their birth county (.177).

Model 2 included a Level 1 predictor to examine whether the amount of government expenditures on early childhood programs was associated with the odds of being competent. In Model 2, maltreated participants whose birth counties were allocated higher amounts of funding for Smart Start and More at Four during the years in which participants were age 0-5 were found to have significantly greater odds of being competent than participants whose birth counties

received lower amounts of funding, $F(2, 202) = 729.020$, $p < .0001$. The estimated odds ratio for large compared to small expenditures on early childhood programs was 2.593, 95% CI [2.466, 2.725], indicating that maltreated participants whose birth counties received relatively large allocations for expenditures on these programs had 2.5 times the odds of being competent compared to maltreated participants whose birth counties received relatively small allocations. The estimated odds ratio for moderate compared to small investments in early childhood programs was 1.810, 95% CI [1.739, 1.883]. Fit statistics for Model 2 indicated improved fit over the unconditional model. The pattern of results was similar for participants in the comparison group but with a slightly smaller estimated odds ratio with respect to large compared to small investments (*OR* = 2.166, 95% CI [2.113, 2.220]).

Model 3 included several Level 1 demographic variables and covariates to control for the effects of gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year (age), and years of education data. County expenditures on Smart Start and More at Four continued to predict competence after controlling for the effects of the covariates included in Model 3. The estimated odds ratio for large compared to small expenditures on Smart Start and More at Four adjusted for covariates was 2.432, 95% CI [2.309, 2.563], indicating that maltreated participants living in counties with relatively large expenditures on early childhood programs had almost 2.5 times the odds of being competent than maltreated participants born in counties with relatively small expenditures on early childhood programs. Maltreated participants living in counties with moderate compared to small expenditures on Smart Start and More at Four had almost twice the odds of being consistently competent (*OR* = 1.730, 95% CI [1.659, 1.803]. The pattern of results was similar for participants in the comparison group. Model fit statistics indicated improved fit over Model 2 with county expenditures as the only Level 1 variable.

3.3.5 Subgroup Analyses Including Movers in School County

Analyses using the full sample assuming that movers lived in their school county indicated that there were no significant differences between the maltreated group and comparison group with respect to the effect of county expenditures on participants' competent functioning over time. However, given the focus of this study on competent functioning among maltreated participants specifically and for the sake of completeness, the effects of county expenditures on Smart Start and More at Four on consistent competence within the maltreated group and the comparison group were analyzed separately to identify any differential impacts across groups.

The Model 1 variance terms for the random effect of county on competence for the full maltreated group and the full comparison group assuming that movers lived in their school county immediately following birth were the same as in the previous analysis (.130 and .187, respectively), as the only differences between the two sets of analyses including movers involved the Level 1 predictor. In Model 2, which added a Level 1 predictor to the unconditional model, maltreated participants whose counties received higher amounts of funding for Smart Start and More at Four during the years in which participants were age 0-5 were found to have significantly greater odds of being competent than participants whose counties received lower amounts of funding, $F(2, 199) = 1{,}058.160$, $p < .0001$.

The estimated odds ratio for large compared to small expenditures on early childhood programs was 3.948, 95% CI [3.719, 4.191], indicating that maltreated participants whose counties received relatively large allocations for expenditures on these programs had 4 times the odds of being consistently competent compared to maltreated participants whose counties received relatively small allocations. The estimated odds ratio for moderate compared to small expenditures on early childhood programs was 2.168, 95% CI [2.075, 2.265]. Fit statistics for Model 2 indicated improved fit over the unconditional model. The pattern of results was similar

for participants in the comparison group but with a smaller estimated odds ratio with respect to large compared to small expenditures (*OR* = 3.199, 95% CI [3.104, 3.296]).

County expenditures on Smart Start and More at Four continued to predict competent functioning after controlling for the effects of the Model 3 covariates. The estimated odds ratio for large compared to small expenditures on Smart Start and More at Four adjusted for covariates was 3.593, 95% CI [3.375, 3.827], indicating that maltreated participants in counties with relatively large expenditures on early childhood programs had 3.5 times the odds of being competent compared to maltreated participants in counties with relatively small expenditures on early childhood programs. The estimated odds ratio for moderate compared to small expenditures on early childhood programs was 2.074, 95% CI [1.982, 2.172]. The pattern of results was similar for participants in the comparison group but with a larger estimated odds ratio with respect to large compared to small expenditures (*OR* = 4.037, 95% CI [3.908, 4.171]). Model fit statistics indicated improved fit over Model 2 with county expenditures as the only Level 1 variable.

3.3.6 Subgroup Analyses Including Nonmovers Only

Analyses using only nonmovers from the full sample indicated that there were no significant differences between the maltreated group and comparison group with respect to the effect of county expenditures on participants' competent functioning over time. However, in light this study's interest in investigating competent functioning among participants in the maltreated group specifically and for the sake of completeness, the effects of county expenditures on Smart Start and More at Four on consistent competence within the maltreated group and the comparison group were analyzed separately to investigate any differential impacts across groups.

The Model 1 variance term for the random effect of county on consistent competence for the nonmover subsample of the maltreated group was .145, indicating that approximately 4% of the total variance in the odds of a participant being competent could be attributed to differences

between counties. The variance term was similar but somewhat larger for the nonmover subsample of the comparison group (.187). In Model 2, maltreated participants whose counties were allocated higher amounts of funding for Smart Start and More at Four during the years in which participants were age 0-5 were found to have significantly greater odds of being competent than participants whose counties were allocated lower amounts of funding, $F(2, 186) = 847.250$, $p < .0001$.

The estimated odds ratio for large compared to small expenditures on early childhood programs was 4.881, 95% CI [4.521, 5.271], indicating that nonmovers in the maltreated group whose counties received relatively large allocations for expenditures on these programs had almost 5 times the odds of being competent compared to nonmovers in the maltreated group whose counties received relatively small allocations. The estimated odds ratio for moderate compared to small investments in early childhood programs was 2.353, 95% CI [2.228, 2.486]. Fit statistics for Model 2 indicated improved fit over the unconditional model. The pattern of results was similar for nonmovers in the comparison group but with a smaller estimated odds ratio with respect to large compared to small investments ($OR = 3.466$, 95% CI [3.343, 3.593]).

In Model 3, county expenditures on Smart Start and More at Four continued to predict consistent competence after controlling for the effects of the covariates. The estimated odds ratio for large compared to small expenditures on Smart Start and More at Four adjusted for covariates was 4.349, 95% CI [4.012, 4.714], indicating that nonmovers in the maltreated group living in counties with relatively large expenditures on early childhood programs had more than 4 times the odds of being competent compared to nonmovers in the maltreated group living in counties with relatively small expenditures on early childhood programs. The estimated odds ratio for moderate compared to small expenditures on early childhood programs was 2.115, 95% CI [2.091, 2.345]. The pattern of results was similar for nonmovers in the comparison group. Model fit

statistics indicated improved fit over Model 2 with county expenditures as the only Level 1 variable.

Table 13: Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four – Maltreated Group, Movers in Birth County)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-1.471*** (.038)	-2.044*** (.048)			-2.698*** (.047)		
Smart Start and More at Four county expenditures			(2, 202)	729.020***		(2, 202)	580.730***
Investment > $2,500		.953*** (.025)			.889*** (.026)		
Investment $1,500 - $2,500		.593*** (.020)			.548*** (.021)		
Investment $500 - $1,500		Reference			Reference		
Gender						(1, 101)	2,405.910***
Female					.705*** (.014)		
Male					Reference		
Race/ethnicity						(3, 290)	570.550***
Black					-.814*** (.020)		
Hispanic					-.543*** (.038)		
Other					-.179*** (.024)		
White					Reference		
SES						(2, 200)	1,885.460***
High SES					1.313*** (.024)		
Middle SES					.765*** (.017)		
Low SES					Reference		
Fetal well-being						(1, 100)	30.360***
Full-term and normal BW					.112*** (.020)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	15.560***
Married					.064*** (.016)		
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.054*** (.002)	(1, 153479)	1,032.890***
Mother's age at birth (GMC)					-.006*** (.001)	(1, 153479)	20.340***
Level 2		Random parameters					
County-level variance	.130 (.020)	.194 (.030)			.118 (.019)		
-2LL (random intercept)	143,952.000	143,193.000			130,756.000		
Chi-square (random intercept)	3,041.700***	3,815.880***			2,407.200***		
-2LL (random effects)	140,528.000	138,956.800			127,985.000		
Pearson chi-square	153,017.900	153,366.800			156,756.800		
Pearson chi-square / DF	1.000	1.000			1.020		
Model fit statistics							
-2LL	140,910.100	139,377.300			128,348.700		
AIC	140,914.100	139,385.300			128,376.700		
BIC	140,919.400	139,395.800			128,413.500		

Note. Sample used in analyses included all participants in maltreated group. Participants who moved from their birth county and lived in another county during most or all education years (movers) were assumed to live in birth county during ages 0-5. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years participants were age 0-5. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.

***All parameter estimates were statistically significant at $p < .001$.

Table 14: Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four – Maltreated Group, Movers in School County)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-1.471*** (.038)	-2.263*** (.059)			-2.904*** (.056)		
Smart Start and More at Four county expenditures			(2, 202)	729.020***		(2, 199)	831.410***
Investment > $2,500		1.373*** (.030)			1.279*** (.032)		
Investment $1,500 - $2,500		.774*** (.022)			.730*** (.023)		
Investment $500 - $1,500		Reference			Reference		
Gender						(1, 101)	2,429.720***
Female					.712*** (.014)		
Male					Reference		
Race/ethnicity						(3, 290)	562.630***
Black					-.810*** (.020)		
Hispanic					-.558*** (.039)		
Other					-.180*** (.024)		
White					Reference		
SES						(2, 200)	1,888.180***
High SES					1.324*** (.024)		
Middle SES					.761*** (.017)		
Low SES					Reference		
Fetal well-being						(1, 100)	30.160***
Full-term and normal BW					.112** (.020)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	17.440***
Married					.067*** (.016)		
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.050*** (.002)	(1, 153479)	849.610***
Mother's age at birth (GMC)					-.007*** (.001)	(1, 153479)	220.490***
Level 2		Random parameters					
County-level variance	.130 (.020)	.300 (.045)			.197 (.031)		
-2LL (random intercept)	143,952.000	143,139.000			130,718***		
Chi-square (random intercept)	3,041.700***	4,497.990***			2,941.520***		
-2LL (random effects)	140,528.000	138,178.300			127,363.300		
Pearson chi-square	153,017.900	153,061.300			156,301.300		
Pearson chi-square / DF	1.000	1.000			1.020		
Model fit statistics							
-2LL	140,910.100	138,641.100			127,776.700		
AIC	140,914.100	138,649.100			127,804.700		
BIC	140,919.400	138,659.600			127,841.500		

Note. Sample used in analyses included all participants in maltreated group. Participants who moved from birth county and lived in another county during most or all education years (movers) were assumed to live in school county during ages 0-5. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years that participants were age 0-5. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.

***All parameter estimates were statistically significant at $p < .001$.

Table 15: Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four – Maltreated Group, Nonmovers)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-1.502*** (.041)	-2.416*** (.069)			-3.018*** (.066)		
Smart Start and More at Four county expenditures			(2, 186)	847.250***		(2, 186)	658.040***
Investment > $2,500		1.585*** (.039)			1.470*** (.041)		
Investment $1,500 - $2,500		.856*** (.028)			.795*** (.029)		
Investment $500 - $1,500		Reference			Reference		
Gender						(1, 100)	1,671.170***
Female					.698*** (.017)		
Male					Reference		
Race/ethnicity						(3, 274)	439.370***
Black					-.842*** (.024)		
Hispanic					-.586*** (.044)		
Other					-.195*** (.029)		
White					Reference		
SES						(2, 193)	1,403.160***
High SES					1.347*** (.029)		
Middle SES					.792*** (.020)		
Low SES					Reference		
Fetal well-being						(1, 100)	20.500***
Full-term and normal BW					.110*** (.024)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	10.840**
Married					.064*** (.019)		$p = .001$
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.054*** (.002)	(1, 113511)	691.040***
Mother's age at birth (GMC)					-.006*** (.002)	(1, 113511)	13.770***
Level 2		Random parameters					
County-level variance	.145 (.024)	.385 (.060)			.258 (.042)		
-2LL (random intercept)	104,376.000	103,753.000			94,200.000		
Chi-square (random intercept)	2,616.310***	3,835.910***			2,516.040***		
-2LL (random effects)	101,417.700	99,482.980			91,296.420		
Pearson chi-square	113,055.300	113,438.900			116,476.500		
Pearson chi-square / DF	1.000	1.000			1.030		
Model fit statistics							
-2LL	101,760.100	99,916.760			91,683.560		
AIC	101,764.100	99,924.760			91,711.560		
BIC	101,769.400	99,935.260			91,748.310		

Note. Sample used in analyses was subsample of participants who remained in same county at birth and during most or all education years (nonmovers) and were assumed to live in birth county during ages 0-5. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years participants were age 0-5. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion. ***All parameter estimates were statistically significant at $p < .001$ except as noted. **$p < .01$.

Table 16: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four - Maltreated Group)

Variable	Include movers in birth county		Include movers in school county		Nonmovers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
Smart Start and More at Four county expenditures						
Investment > $2,500 vs. $500 - $1,500	2.593	[2.466, 2.725]	3.948	[3.719, 4.191]	4.881	[4.521, 5.271]
Investment $1,500 - $2,500 vs. $500 - $1,500	1.810	[1.739, 1.883]	2.168	[2.075, 2.265]	2.353	[2.228, 2.486]
			Model 3			
Smart Start and More at Four county expenditures						
Investment > $2,500 vs. $500 - $1,500	2.432	[2.309, 2.563]	3.593	[3.375, 3.827]	4.349	[4.012, 4.714]
Investment $1,500 - $2,500 vs. $500 - $1,500	1.730	[1.659, 1.803]	2.074	[1.982, 2.172]	2.115	[2.091, 2.345]
Gender						
Female vs. male	2.025	[1.968, 2.083]	2.036	[1.978, 2.095]	2.009	[1.942, 2.078]
Race/ethnicity						
White vs. Black	2.257 *(IOR)*	[2.169, 2.347]	2.247 *(IOR)*	[2.160, 2.336]	2.320 *(IOR)*	[2.212, 2.433]
White vs. Hispanic	1.721 *(IOR)*	[1.595, 1.855]	1.748 *(IOR)*	[1.621, 1.883]	1.795 *(IOR)*	[1.647, 1.961]
White vs. Other	1.196 *(IOR)*	[1.140, 1.253]	1.198 *(IOR)*	[1.142, 1.255]	1.215 *(IOR)*	[1.148, 1.287]
SES						
High SES vs. low SES	3.718	[3.546, 3.899]	3.759	[3.584, 3.942]	3.844	[3.633, 4.067]
Middle SES vs. low SES	2.148	[2.079, 2.221]	2.141	[2.071, 2.213]	2.207	[2.122, 2.296]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.119	[1.075, 1.165]	1.119	[1.074, 1.165]	1.116	[1.063, 1.171]
Mother's marital status						
Married vs. not married	1.066	[1.032, 1.100]	1.070	[1.036, 1.105]	1.066	[1.026, 1.107]
Birth year (GMC, RC) X years of education data (GMC)	1.056	[1.052, 1.059]	1.051	[1.048, 1.055]	1.056	[1.051, 1.060]
Mother's age at birth (GMC)	1.006 *(IOR)*	[1.004, 1.009]	1.007 *(IOR)*	[1.004, 1.009]	1.006 *(IOR)*	[1.003, 1.009]

Note. Samples used in analyses were: full sample of maltreated participants assuming movers lived in birth county during ages 0-5, full sample of maltreated participants assuming movers lived in the county in which they attended school for most or all years of available education data during ages 0-5, and subsample of maltreated participants who remained in same county at birth and during most or all education years (nonmovers) and were assumed to live in birth county during ages 0-5. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years that participants were age 0-5. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

Table 17: Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four – Comparison Group, Movers in Birth County)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-.618*** (.042)	-1.127*** (.048)			-2.386*** (.043)		
Smart Start and More at Four county expenditures			(2, 202)	2,534.240***		(2, 202)	3,179.200***
Investment > $2,500		.773*** (.012)			.959*** (.014)		
Investment $1,500 - $2,500		.572*** (.009)			.678*** (.010)		
Investment $500 - $1,500		Reference			Reference		
Gender						(1, 100)	8,792.740***
Female					.656*** (.007)		
Male					Reference		
Race/ethnicity						(3, 295)	2,468.890***
Black					-.867*** (.011)		
Hispanic					-.797*** (.017)		
Other					-.166*** (.013)		
White					Reference		
SES						(2, 202)	7,994.630***
High SES					1.363*** (.011)		
Middle SES					.710*** (.011)		
Low SES					Reference		
Fetal well-being						(1, 100)	158.650***
Full-term and normal BW					.139*** (.011)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	181.800***
Married					.136*** (.010)		
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.047*** (.001)	(1, 454871)	1,659.860***
Mother's age at birth (GMC)					.009*** (.001)	(1, 454871)	174.300***
Level 2		Random parameters					
County-level variance	.177 (.026)	.227 (.033)			.154 (.022)		
-2LL (random intercept)	591,709.000	590,054.000			520,732.000		
Chi-square (random intercept)	28,681.400***	3,2270.400***			28,721.000***		
-2LL (random effects)	562,472.500	557,203.900			491,485.900		
Pearson chi-square	454,562.100	455,751.100			470,738.600		
Pearson chi-square / DF	1.000	1.000			1.030		
Model fit statistics							
-2LL	563,027.800	557,783.400			492,011.300		
AIC	563,031.800	557,791.400			492,039.300		
BIC	563,037.100	557,801.900			492,076.100		

Note. Sample used in analyses was full comparison group. Participants who moved from birth county and lived in another county during most or all education years (movers) were assumed to have lived in birth county during ages 0-5. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years that participants were age 0-5. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.
***All parameter estimates were statistically significant at $p < .001$.

Table 18: Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four – Comparison Group, Movers in School County)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-.618*** (.042)	-1.314*** (.057)			-2.623*** (.053)		
Smart Start and More at Four county expenditures			(2, 199)	3,559.800***		(2, 199)	4,405.400***
Investment > $2,500		1.163*** (.015)			1.396*** (.017)		
Investment $1,500 - $2,500		.713*** (.010)			.873*** (.011)		
Investment $500 - $1,500		Reference			Reference		
Gender						(1, 100)	8,840.340***
Female					.660*** (.007)		
Male					Reference		
Race/ethnicity						(3, 295)	2,465.800***
Black					-.862*** (.011)		
Hispanic					-.822*** (.017)		
Other					-.160*** (.013)		
White					Reference		
SES						(2, 202)	8,121.760***
High SES					1.378*** (.011)		
Middle SES					.711*** (.011)		
Low SES					Reference		
Fetal well-being						(1, 100)	167.510***
Full-term and normal BW					.143** (.011)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 101)	194.610***
Married					.142*** (.010)		
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.045*** (.001)	1, 454871	1,480.240***
Mother's age at birth (GMC)					.009*** (.001)	1, 454871	152.650***
Level 2		Random parameters					
County-level variance	.177 (.026)	.320 (.047)			.197 (.031)		
-2LL (random intercept)	591,709.000	590,266.000			130,718***		
Chi-square (random intercept)	28,681.400***	34,686.700***			2,941.520***		
-2LL (random effects)	562,472.500	554,966.600			127,363.300		
Pearson chi-square	454,562.100	454,937.700			156,301.300		
Pearson chi-square / DF	1.000	1.000			1.020		
Model fit statistics							
-2LL	563,027.800	555,579.700			127,776.700		
AIC	563,031.800	555,587.700			127,804.700		
BIC	563,037.100	555,598.200			127,841.500		

Note. Sample used in analyses was full comparison group. Participants who moved from birth county and lived in another county during most or all education years (movers) were assumed to have lived in their school county during ages 0-5. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years participants were age 0-5. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.

***All parameter estimates were statistically significant at $p < .001$.

Table 19: Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four – Comparison Group, Nonmovers)

Parameter	Model 1	Model 2			Model 3		
	Estimate	Estimate	DF	*F*	Estimate	DF	*F*
Level 1		Fixed effects					
Intercept	-.632*** (.044)	-1.375*** (.061)			-2.685*** (.059)		
Smart Start and More at Four county expenditures			(2, 186)	2,745.860***		(2, 186)	3,529.080***
Investment > $2,500		1.243*** (.018)			1.522*** (.020)		
Investment $1,500 - $2,500		.728*** (.012)			.910*** (.012)		
Investment $500 - $1,500		Reference			Reference		
Gender						(1, 100)	6,529.330***
Female					.656*** (.008)		
Male					Reference		
Race/ethnicity						(3, 282)	1,980.850***
Black					-.897*** (.013)		
Hispanic					-.859*** (.019)		
Other					-.179*** (.016)		
White					Reference		
SES						(2, 200)	5,930.090***
High SES					1.372*** (.013)		
Middle SES					.705*** (.012)		
Low SES					Reference		
Fetal well-being						(1, 98)	122.080***
Full-term and normal BW					.142*** (.013)		
Premature or low BW					Reference		
Mother's marital status at birth						(1, 100)	124.640***
Married					.134*** (.012)		
Not married					Reference		
Birth year (GMC, RC) X years of education data (GMC)					.047*** (.001)	(1, 346458)	1,221.200***
Mother's age at birth (GMC)					.010*** (.001)	(1, 346458)	146.140***
Level 2		Random parameters					
County-level variance	.187 (.028)	.349 (.052)			.286 (.042)		
-2LL (random intercept)	448,971.000	447,884.000			394,664.000		
Chi-square (random intercept)	24,745.4***	29,406.700***			27,395.300***		
-2LL (random effects)	423,712.100	417,904.500			366,731.300		
Pearson chi-square	346,138.700	346,350.100			358,546.700		
Pearson chi-square / DF	1.000	1.000			1.030		
Model fit statistics							
-2LL	424,225.200	418,477.000			367,268.800		
AIC	424,229.200	418,485.000			367,296.800		
BIC	424,234.400	418,495.500			367,333.600		

Note. Sample used in analyses was subsample of participants in comparison sample who remained in same county at birth and during most or all education years (nonmovers) and were assumed to live in birth county during ages 0-5. Model 1 is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start and More At Four on children age 0-5 within the county during years participants were age 0-5. Model 3 includes several Level 1 covariates. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; -2LL = -2 log likelihood; DF = degrees of freedom; AIC = Akaike's Information Criterion; BIC = Schwartz's Bayesian Information Criterion.

***All parameter estimates were statistically significant at $p < .001$.

Table 20: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start and More at Four - Comparison Group)

Variable	Include movers in birth county		Include movers in school county		Nonmovers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
Smart Start and More at Four county expenditures						
Investment > $2,500 vs. $500 - $1,500	2.166	[2.113, 2.220]	3.199	[3.104, 3.296]	3.466	[3.343, 3.593]
Investment $1,500 - $2,500 vs. $500 - $1,500	1.772	[1.742, 1.804]	2.039	[2.000, 2.079]	2.070	[2.024, 2.118]
			Model 3			
Smart Start and More at Four county expenditures						
Investment > $2,500 vs. $500 - $1,500	2.608	[2.539, 2.679]	4.037	[3.908, 4.171]	4.582	[4.406, 4.765]
Investment $1,500 - $2,500 vs. $500 - $1,500	1.971	[1.934, 2.008]	2.394	[2.344, 2.444]	2.485	[2.424, 2.548]
Gender						
Female vs. male	1.927	[1.900, 1.953]	1.935	[1.908, 1.962]	1.926	[1.896, 1.958]
Race/ethnicity						
White vs. Black	2.381 *(IOR)*	[2.331, 2.433]	2.370 *(IOR)*	[2.320, 2.421]	2.451 *(IOR)*	[2.392, 2.513]
White vs. Hispanic	2.217 *(IOR)*	[2.146, 2.294]	2.278 *(IOR)*	[2.203, 2.353]	2.358 *(IOR)*	[2.273, 2.451]
White vs. Other	1.181 *(IOR)*	[1.149, 1.212]	1.174 *(IOR)*	[1.143, 1.205]	1.196 *(IOR)*	[1.160, 1.233]
SES						
High SES vs. low SES	3.908	[3.825, 3.993]	3.968	[3.883, 4.055]	3.943	[3.845, 4.044]
Middle SES vs. low SES	2.035	[1.993, 2.078]	2.035	[1.993, 2.079]	2.023	[1.974, 2.073]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.149	[1.124, 1.175]	1.154	[1.129, 1.180]	1.152	[1.123, 1.182]
Mother's marital status						
Married vs. not married	1.146	[1.123, 1.169]	1.152	[1.129, 1.175]	1.143	[1.116, 1.170]
Birth year (GMC, RC) X years of education data (GMC)	1.048	[1.046, 1.051]	1.046	[1.044, 1.048]	1.048	[1.045, 1.051]
Mother's age at birth (GMC)	1.009	[1.008, 1.011]	1.009	[1.007, 1.010]	1.010	[1.008, 1.011]

Note. Samples used in analyses were: full comparison group assuming movers lived in their birth county during ages 0-5, full comparison group assuming movers lived in the county in which they attended school for most or all of the years of available education data during ages 0-5, and subsample of comparison group who remained in same county at birth and during most or all education years (nonmovers) and were assumed to have lived in birth county during ages 0-5. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on the Smart Start and More At Four programs on children age 0-5 within the county during the years that participants were age 0-5. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

3.3.7 Multilevel Analyses of the Effect of Expenditures on Smart Start and More at Four Programs on Patterns of Competence

In addition to the primary multilevel analyses investigating the effect of county expenditures on Smart Start and More at Four on participants' consistently competent functioning across time, supplemental multilevel analyses examined the effects of the relative levels of expenditures on these early childhood programs on patterns of competence among participants. Multilevel logistic regression models were used to investigate effects on consistent competence across time within specific domains of functioning, including academic performance, special education, and behavioral domains. Multilevel regression analyses were used to examine effects on the percentage of years in which participants demonstrated competence across all domains.

After controlling for gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data as well as maltreatment status, relative expenditures on Smart Start and More at Four within participants' county when they were age 0-5 predicted competence across time in the domain of academic functioning, $F(2, 202) = 3{,}862.660$, $p < .0001$. Participants who lived in counties with relatively large expenditures on early childhood programs had almost 3 times the odds of being competent in the academic performance domain than participants who lived in counties with relatively small expenditures on early childhood programs ($OR = 2.798$, 95% CI [2.720, 2.854]), and participants who lived in counties with moderate levels of expenditures had about twice the odds of being competent in the academic performance domain than participants who lived in counties with relatively small expenditures ($OR = 1.938$, 95% CI [1.903, 1.974]).

The effect of maltreatment status on academic performance was significant after controlling for the effects of Smart Start and More at Four as well as the covariates, $F(1, 101) = 709.000$, $p < .0001$, such that nonmaltreated participants had somewhat greater odds of functioning competently in the academic performance domain across time than maltreated

participants (*OR* = 1.227, 95% CI [1.208, 1.246]). In addition, the interaction between county expenditures and maltreatment status was significant but small, $F(2, 202) = 42.420$, $p < .0001$, indicating that there might be differences in the effects of expenditures on Smart Start and More at Four on competence within the academic performance domain across groups. The small size of the effect suggested that group differences were not likely to be substantial.

In analyses run separately by group, the effects of expenditures on Smart Start and More at Four on competence within the academic performance domain were similar across groups but slightly larger in the comparison group than in the maltreated group. For the comparison group, the estimated odds ratios for large compared to small expenditures and for moderate compared to small expenditures were 2.910, 95% CI [2.836, 2.986], and 2.098, 95% CI [2.061, 2.137], respectively. For the maltreated group, the estimated odds ratios for large compared to small expenditures and for moderate compared to small expenditures were 2.538, 95% CI [2.430, 2.650], and 1.749, 95% CI [1.692, 1.809], respectively.

Results of analyses examining effects of expenditures on Smart Start and More at Four on consistently competent functioning within the behavioral domain were similar to those reported for the academic performance domain. The effect of maltreatment status on behavioral functioning was significant after controlling for the effects of county expenditures and the covariates, $F(1, 101) = 5{,}062.480$, $p < .0001$, such that nonmaltreated participants had greater odds of functioning competently in the behavioral domain across time than maltreated participants (*OR* = 1.728, 95% CI [1.702, 1.755]). The interaction between county expenditures and maltreatment status was not significant, $F(2, 202) = 2.330$, $p = .100$, indicating that there were no meaningful differences in effects across groups.

The positive effects of expenditures on Smart Start and More at Four on consistently competent functioning overall as well as within the academic performance and behavioral

functioning domains across time did not hold with respect to the special education domain. There was a slight negative effect of large compared to small expenditures (*IOR* = 1.055, 95% CI [1.029, 1.081]) and moderate compared to small expenditures (*IOR* = 1.047, 95% CI [1.028, 1.067]). However, this effect was very small ($F(2, 202) = 12.850$, $p < .0001$), indicating that there was likely not a meaningful difference across levels of expenditures on Smart Start and More at Four on competence within the special education domain. In subgroup analyses, the effect of government expenditures on competence within the special education domain was not significant for the maltreated group, $F(2, 202) = 2.850$, $p = .060$, and there was a very small negative effect for the comparison group, $F(2, 202) = 6.980$, $p = .001$. In the full sample, the effect of maltreatment status on competence in the special education domain was significant after controlling for the effects of Smart Start and More at Four as well as the covariates, $F(1, 101) = 1{,}042.390$, $p < .0001$. However, the interaction between county expenditures and maltreatment status was not significant, $F(2, 202) = 1.320$, $p = .268$, indicating no meaningful differences in the effect of expenditures on Smart Start and More at Four across groups.

In multilevel regression analyses including expenditures on Smart Start and More at Four as a Level 1 predictor as well as the covariates, living in a county with relatively large expenditures on early childhood programs predicted being competent across domains for a greater percentage of years, $F(2, 202) = 5{,}615.670$, $p < .0001$. The effect of maltreatment status on the percentage of years in which participants were competent was significant after controlling for the effects of Smart Start and More at Four as well as the covariates, $F(1, 101) = 11{,}777.600$, $p < .0001$. In addition, the interaction between county expenditures and maltreatment status was significant but small, $F(2, 202) = 16.370$, $p < .0001$, indicating that there might be differences in the effects of county expenditures across groups. The small size of the effect suggested that group differences were not likely to be substantial. In analyses by subgroup, relative levels of

expenditures on Smart Start and More at Four predicted the percentage of years in which participants were competent for both the maltreated group, $F(1, 101) = 1{,}466.140$, $p < .0001$, and the comparison group, $F(1, 101) = 5{,}307.230$, $p < .0001$.

3.3.8 Supplemental Analyses on the Separate Effects of Smart Start and More at Four Programs

Though the primary research question for these sets of analyses focused on the effects of total county expenditures for early childhood programs on consistently competent functioning, a logical follow-up question was the relative benefit associated with expenditures for each program. Supplemental analyses examined whether Smart Start and More at Four each had an independent influence on competence among participants and also investigated whether the pattern of results was similar for maltreated and nonmaltreated children.

3.3.8.1 Separate effects of More at Four on consistent competence

Total per capita expenditures on More at Four were categorized as none ($0, approximately 19% of the maltreated group and 16% of the comparison group), small (between $0 and $200, approximately 15% of each group), moderate (between $200 and $600, approximately 43% of the maltreated group and 47% of the comparison group), and large (greater than $600, approximately 23% of the maltreated group and 22% of the comparison group). These relative percentages were consistent across the three different subsamples of movers and nonmovers within each group used in these analyses.

Prior to conducting multilevel analyses, the relative percentages of competent and non-competent participants at each level of county expenditures on More at Four were calculated. Among participants living in counties with no expenditures on More at Four, 19% were competent, whereas among participants living in counties with large expenditures on More at Four, 41% were competent. Twenty-six percent (26%) of participants living in counties with

relatively small expenditures and 33% of participants living in counties with relatively moderate expenditures on More at Four were competent. Figure 4 depicts the interaction of Smart Start and More at Four expenditures by maltreatment status on participants' consistent competence without adjusting for any covariates. Relative levels of government expenditures on More at Four predicted consistent competence ($\chi^2(3) = 14{,}824.862$, $p < .0001$).

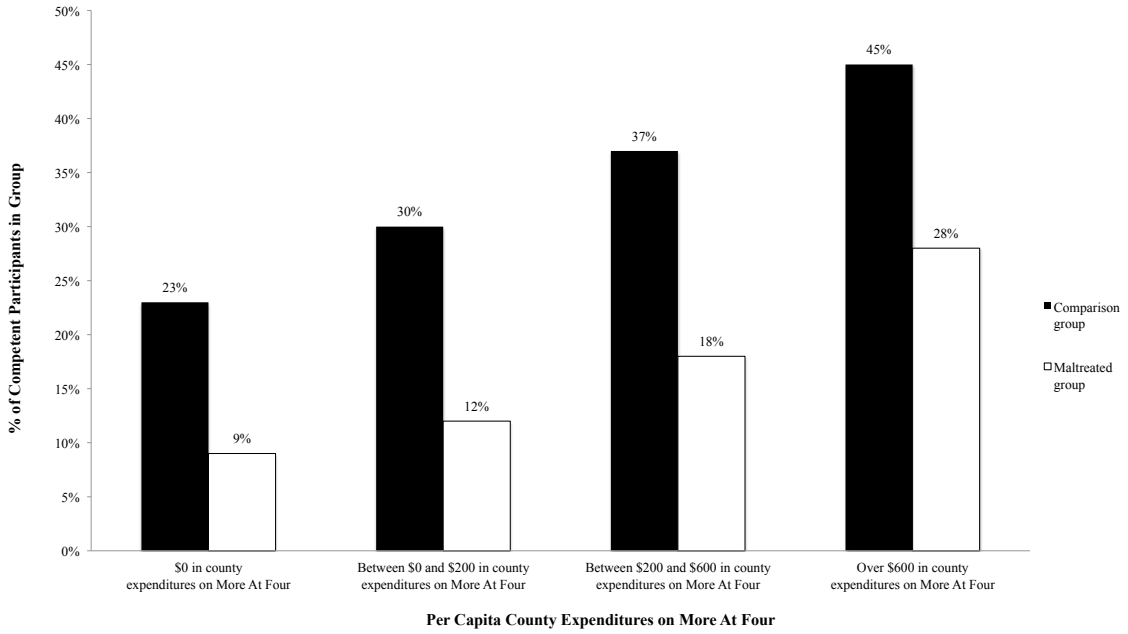

Figure 4: Consistent Competence by More at Four Expenditures and Maltreatment Status

The overall effect of government expenditures on More at Four on participants' competence was examined using multilevel logistic regression. The first set of analyses included all participants assuming movers remained in their birth county during the ages of 0-5 and thus received the benefit of expenditures on More at Four in their birth county. In a model including levels of expenditures on More at Four as a Level 1 predictor without adjusting for covariates, the estimated odds ratio for large expenditures compared to no expenditures on More at Four was 3.630, 95% CI [3.557, 3.705], the estimated odds ratio for moderate expenditures compared to no expenditures on More at Four was 2.180, 95% CI [2.140, 2.220], and the estimated odds ratio for

small expenditures compared to no expenditures on More at Four was 1.337, 95% CI [1.307, 1.368].

In a model adding maltreatment status, the interaction of maltreatment status and expenditures on More at Four as additional Level 1 predictors, and the full set of covariates used in the previous analyses, the estimated odds ratios were similar or somewhat larger. Participants who lived in counties with relatively large expenditures on More at Four during the year in which participants were 4 years of age had 4.5 times the odds of being consistently competent as participants who lived in counties with no expenditures on More at Four (*OR* = 4.505, 95% CI [4.379, 4.635]). Participants who lived in counties with moderate expenditures on More at Four had 2.5 times the odds of being competent as maltreated participants who lived in counties with no expenditures on More at Four (*OR* = 2.559, 95% CI [2.493, 2.628]). Participants in the maltreated group who lived in counties with relatively small expenditures on More at Four had significantly greater odds of being competent than participants in the maltreated group who lived in counties with no expenditures on More at Four (*OR* = 1.380, 95% CI [1.336, 1.426]). In addition, after controlling for the effects of More at Four expenditures and covariates, participants in the comparison group had 1.5 times the odds of being competent as participants in the maltreated group (*OR* = 1.476, 95% CI [1.447, 1.505]). The interaction of More at Four expenditures and maltreatment status was significant, suggesting that More at Four expenditures might have a differential effect on competent functioning on maltreated and nonmaltreated participants, but the interaction effect was quite small, $F(3, 299) = 3.34$, $p = .020$.

The results were similar for analyses including all participants assuming that the movers moved to their school county after birth. In a model including expenditures on More at Four, maltreatment status, and the interaction of maltreatment status and expenditures on More at Four as Level 1 predictors of consistent competence and adjusting for covariates, the estimated odds

ratios were as follows: *OR* = 4.930, 95% CI [4.790, 5.074], for large compared to no expenditures, *OR* = 2.545, 95% CI [2.479, 2.613], for moderate compared to no expenditures, and *OR* = 1.316, 95% CI [1.273, 1.361], for small compared to no expenditures. After controlling for the effects of More at Four expenditures and covariates, participants in the comparison group had 1.5 times the odds of being consistently competent as participants in the maltreated group (*OR* = 1.482, 95% CI [1.453, 1.512]). The interaction of More at Four expenditures and maltreatment status was not significant, $F(3, 292) = 1.25$, $p = .291$, indicating that the effects of More at Four expenditures on competence were similar across the maltreated and comparison groups.

A similar pattern of results was found in analyses using only the nonmover participants in the full sample. In a model with the full set of covariates used in the previous analyses as well as expenditures on More at Four, maltreatment status, and the interaction of maltreatment status and expenditures on More at Four as Level 1 predictors, the estimated odds ratios were as follows: *OR* = 5.135, 95% CI [4.961, 5.316], for large compared to no expenditures, *OR* = 2.597, 95% CI [2.516, 2.682], for moderate compared to no expenditures, and *OR* = 1.310, 95% CI [1.258, 1.363], for small compared to no expenditures. Participants in the comparison group had 1.5 times the odds of being consistently competent as participants in the maltreated group (*OR* = 1.487, 95% CI [1.452, 1.523]). The interaction of More at Four expenditures and maltreatment status was not significant, $F(3, 265) = 2.22$, $p = .086$, indicating that there were no significant differences in the effect of government expenditures on More at Four on consistent competence among nonmovers in the maltreated and comparison groups.

Though only one out of the three foregoing analyses using the full sample indicated that the effects of More at Four expenditures might be different across groups, subsequent analyses examined the effects of expenditures on More at Four within the maltreated group and comparison group separately. In the maltreated group, the estimated odds ratios without

controlling for covariates were 4.681, 95% CI [4.455, 4.918], for large compared to no expenditures on More at Four, 2.348, 95% CI [2.242, 2.460], for moderately large compared to no expenditures, and 1.299, 95% CI [1.226, 1.377], for small compared to no expenditures, and these effects continued to hold after adjusting for covariates. Maltreated participants who lived in counties with relatively large expenditures on More at Four during the year in which participants were 4 years of age had 4.5 times the odds of being consistently competent as maltreated participants who lived in counties with no expenditures on More at Four (*OR* = 4.569, 95% CI [4.355, 4.816]). Maltreated participants who lived in counties with moderate expenditures on More at Four had almost 2.5 times the odds of being competent as maltreated participants who lived in counties with no expenditures on More at Four (*OR* = 2.417, 95% CI [2.304, 2.536]). Participants in the maltreated group who lived in counties with relatively small expenditures on More at Four had significantly greater odds of being competent than participants in the maltreated group who lived in counties with no expenditures on More at Four (*OR* = 1.321, 95% CI [1.244, 1.402]). The pattern of results for the effects of expenditures on More at Four on maltreated participants' consistently competent functioning was similar across analyses including all maltreated participants assuming that the movers moved to their school county after birth as well as the nonmover subsample of the maltreated group.

The effect of county expenditures for More at Four on consistent competence also was examined for participants in the comparison group to determine whether there was any differential impact on maltreated participants. Effects of More at Four expenditures on consistent competence were similar for participants in the comparison group as those reported for participants in the maltreated group. In analyses using participants in the comparison group including movers assuming that they continued to live in their birth county, participants who lived in counties with relatively large expenditures on More at Four had 4.5 times the odds of

being consistently competent as maltreated participants who lived in counties with no expenditures on More at Four (*OR* = 4.531, 95% CI [4.419, 4.646]). Participants in the comparison group who lived in counties with moderate expenditures on More at Four had more than 2.5 times the odds of being competent as maltreated participants who lived in counties with no expenditures on More at Four (*OR* = 2.671, 95% CI [2.612, 2.731]). Participants in the comparison group who lived in counties with relatively small expenditures on More at Four had significantly greater odds of being competent than participants who lived in counties with no expenditures on More at Four (*OR* = 1.430, 95% CI [1.391, 1.469]). Results were substantially similar in analyses including all participants in the comparison group assuming that the movers moved to their school county after birth as well as analyses using only the nonmover subsample of the comparison group.

Table 21 describes the odds ratios and 95% confidence intervals for the fixed effects in each of three fitted models reporting the effects of More at Four expenditures on competence for the full maltreated group assuming movers remained in their birth county in the relevant years, the full maltreated group assuming movers moved to their school county immediately after birth, and the subsample of nonmovers in the maltreated group, respectively. Table 22 describes the odds ratios and 95% confidence intervals for each fixed effect in three fitted models for the full sample of participants in the comparison group assuming movers remained in their birth county in the relevant years, the full sample of participants in the comparison group assuming movers moved to their school county immediately after their birth, and the subsample of nonmovers in the comparison group, respectively.

Table 21: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Consistent Competence for Children Nested Within School County (More at Four - Maltreated Group)

Variable	Include movers in birth county		Include movers in school county		Nonmovers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
More at Four county expenditures						
Investment > $600 vs. $0	4.681	[4.455, 4.918]	5.116	[4.864, 5.381]	5.524	[5.194, 5.875]
Investment $200 - $600 vs. $0	2.348	[2.242, 2.460]	2.319	[2.213, 2.430]	2.365	[2.234, 2.503]
Investment $100 - $200 vs. $0	1.299	[1.226, 1.377]	1.230	[1.160, 1.305]	1.187	[1.105, 1.275]
			Model 3			
More at Four county expenditures						
Investment > $600 vs. $0	4.569	[4.335, 4.816]	5.013	[4.749, 5.292]	5.345	[5.005, 5.709]
Investment $200 - $600 vs. $0	2.417	[2.304, 2.536]	2.410	[2.296, 2.529]	2.449	[2.310, 2.598]
Investment $100 - $200 vs. $0	1.321	[1.244, 1.402]	1.264	[1.189, 1.343]	1.224	[1.137, 1.318]
Gender						
Female vs. male	2.061	[2.002, 2.121]	2.064	[2.005, 2.124]	2.035	[1.966, 2.106]
Race/ethnicity						
White vs. Black	2.273 *(IOR)*	[2.188, 2.370]	2.262 *(IOR)*	[2.174, 2.358]	2.342 *(IOR)*	[2.237, 2.457]
White vs. Hispanic	1.845 *(IOR)*	[1.709, 1.992]	1.866 *(IOR)*	[1.727, 2.012]	1.927 *(IOR)*	[1.763, 2.105]
White vs. Other	1.229 *(IOR)*	[1.171, 1.289]	1.229 *(IOR)*	[1.172, 1.289]	1.245 *(IOR)*	[1.178, 1.321]
SES						
High SES vs. low SES	3.792	[3.614, 3.979]	3.787	[3.609, 3.974]	3.870	[3.655, 4.098]
Middle SES vs. low SES	2.101	[2.032, 2.173]	2.096	[2.027, 2.168]	2.160	[2.076, 2.247]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.130	[1.085, 1.177]	1.134	[1.088, 1.181]	1.129	[1.076, 1.185]
Mother's marital status						
Married vs. not married	1.099	[1.064, 1.135]	1.104	[1.069, 1.140]	1.097	[1.055, 1.140]
Birth year (GMC, RC) X years of education data (GMC)	1.032	[1.029, 1.036]	1.029	[1.025, 1.032]	1.033	[1.029, 1.037]
Mother's age at birth (GMC)	1.007 *(IOR)*	[1.005, 1.010]	1.008 *(IOR)*	[1.005, 1.010]	1.007 *(IOR)*	[1.004, 1.010]

Note. Samples used in analyses were: full maltreated sample assuming movers lived in their birth county during ages 0-5, full maltreated sample assuming movers lived in the county in which they attended school for most or all of the years of available education data during ages 0-5, and subsample of maltreated participants who remained in same county at birth and during most or all education years (nonmovers) and were assumed to have lived in their birth county during ages 0-5. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on More at Four on 4-year-old children within the county during the year participants were 4 years old. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

Table 22: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Consistent Competence for Children Nested Within School County (More at Four - Comparison Group)

Variable	Include movers in birth county		Include movers in school county		Nonmovers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
More at Four county expenditures						
Investment > $600 vs. $0	3.426	[3.348, 3.505]	3.749	[3.663, 3.838]	3.765	[3.663, 3.869]
Investment $200 - $600 vs. $0	2.135	[2.092, 2.179]	2.086	[2.044, 2.129]	2.075	[2.026, 2.126]
Investment $100 - $200 vs. $0	1.337	[1.304, 1.371]	1.243	[1.211, 1.275]	1.248	[1.211, 1.286]
			Model 3			
More at Four county expenditures						
Investment > $600 vs. $0	4.531	[4.419, 4.646]	4.963	[4.838, 5.091]	5.139	[4.987, 5.295]
Investment $200 - $600 vs. $0	2.671	[2.612, 2.731]	2.637	[2.579, 2.697]	2.700	[2.629, 2.772]
Investment $100 - $200 vs. $0	1.430	[1.391, 1.469]	1.360	[1.323, 1.399]	1.377	[1.333, 1.423]
Gender						
Female vs. male	1.960	[1.933, 1.988]	1.965	[1.938, 1.993]	1.955	[1.924, 1.987]
Race/ethnicity						
White vs. Black	2.410 *(IOR)*	[2.358, 2.463]	2.398 *(IOR)*	[2.347, 2.451]	2.481 *(IOR)*	[2.421, 2.545]
White vs. Hispanic	2.475 *(IOR)*	[2.392, 2.558]	2.519 *(IOR)*	[2.439, 2.604]	2.617 *(IOR)*	[2.519, 2.717]
White vs. Other	1.215 *(IOR)*	[1.183, 1.248]	1.220 *(IOR)*	[1.188, 1.252]	1.244 *(IOR)*	[1.205, 1.282]
SES						
High SES vs. low SES	4.058	[3.970, 4.147]	4.067	[3.979, 4.157]	4.033	[3.931, 4.137]
Middle SES vs. low SES	2.025	[1.982, 2.068]	2.023	[1.981, 2.067]	2.008	[1.958, 2.058]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.168	[1.142, 1.194]	1.169	[1.144, 1.196]	1.164	[1.134, 1.194]
Mother's marital status						
Married vs. not married	1.189	[1.165, 1.213]	1.192	[1.168, 1.216]	1.183	[1.155, 1.212]
Birth year (GMC, RC) X years of education data (GMC)	1.049	[1.047, 1.052]	1.045	[1.043, 1.048]	1.05	[1.047, 1.053]
Mother's age at birth (GMC)	1.008	[1.006, 1.009]	1.007	[1.006, 1.008]	1.008	[1.007, 1.010]

Note. Samples used in analyses were: full comparison group assuming movers lived in their birth county during ages 0-5, full comparison group assuming movers lived in the county in which they attended school for most or all of the years of available education data during ages 0-5, and subsample of comparison group who remained in same county at birth and during most or all education years (nonmovers) and were assumed to have lived in their birth county during ages 0-5. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on More at Four on 4-year-old children within the county during the year participants were 4 years old. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

3.3.8.2 Separate effects of More at Four on patterns of competence

In addition to the supplemental multilevel analyses investigating the separate effects of county expenditures on More at Four on participants' consistently competent functioning across time, additional multilevel analyses examined the effects of the relative levels of expenditures on More at Four on patterns of competence among participants. Multilevel logistic regression models were used to investigate effects on consistent competence across time within specific domains of functioning, including academic performance, special education, and behavioral functioning domains. Multilevel regression analyses were used to examine effects on the percentage of years in which participants demonstrated competence across all domains.

After controlling for gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data as well as maltreatment status, relative expenditures on More at Four within participants' county when they were age 4 predicted competence across time in the domain of academic functioning, $F(2, 202) = 7{,}194.870$, $p < .0001$. Participants who lived in counties with relatively large expenditures on More at Four had almost 5 times the odds of being competent in the academic performance domain than participants who lived in counties with no expenditures on More at Four ($OR = 4.761$, 95% CI [4.652, 4.872]), and participants who lived in counties with moderate levels of expenditures had 2.5 times the odds of being competent in the academic performance domain than participants who lived in counties with no expenditures ($OR = 2.490$, 95% CI [2.439, 2.541]). Participants who lived in counties with small levels of expenditures on More at Four had 1.25 times the odds of being competent in the academic performance domain than participants who lived in counties with no More at Four expenditures ($OR = 1.240$, 95% CI [1.209, 1.272]).

The effect of maltreatment status on academic performance remained significant after controlling for the effect of More at Four as well as the covariates, $F(1, 101) = 497.220$, $p < .0001$, such that nonmaltreated participants had somewhat greater odds of functioning competently in the academic performance domain across time than maltreated participants ($OR = 1.198$, 95% CI [1.179, 1.217]). In addition, the interaction between More at Four expenditures and maltreatment status was significant but small, $F(2, 299) = 53.680$, $p < .0001$, indicating that there might be differences in the effects of expenditures on More at Four on competence within the academic performance domain across groups. The small size of the effect suggested that group differences were not likely to be substantial.

In analyses run separately by group, the effects of relative expenditures on More at Four on competence within the academic performance domain were similar across groups but slightly larger in the comparison group than in the maltreated group. For the comparison group, the estimated odds ratios for large compared to no More at Four expenditures, for moderate compared to no More at Four expenditures, and for small compared to no More at Four expenditures were 5.167, 95% CI [5.046, 5.290], 2.876, 95% CI [2.819, 2.935], and 1.325, 95% CI [1.293, 1.358], respectively. For the maltreated group, the estimated odds ratios for large compared to no expenditures, for moderate compared to no expenditures, and for small compared to no expenditures on More at Four were 4.522, 95% CI [4.340, 4.711], 2.216, 95% CI [2.138, 2.296], and 1.181, 95% CI [1.130, 1.235], respectively.

Results for analyses examining effects of expenditures on More at Four on consistently competent functioning within the behavioral domain were similar to those reported for the academic performance domain. The effect of maltreatment status on behavioral functioning was significant after controlling for the effects of More at Four expenditures and the covariates, $F(1, 101) = 5{,}047.290$, $p < .0001$, such that nonmaltreated participants had greater odds of functioning

competently in the behavioral domain across time than maltreated participants (*OR* = 1.775, 95% CI [1.747, 1.804]). The interaction between More at Four expenditures and maltreatment status was significant, $F(3, 299) = 45.760$, $p < .0001$, indicating potential differences across groups in the effects of More at Four expenditures on competence within the behavioral functioning domain. The small size of the effect suggested that any group differences were likely not substantial.

In analyses run separately by group, the effects of relative expenditures on More at Four on competence within the behavioral functioning domain were similar across groups but slightly larger in the maltreated group than in the comparison group. For the maltreated group, the estimated odds ratios for large compared to no expenditures, for moderate compared to no expenditures, and for small compared to no expenditures were 5.610, 95% CI [5.392], 2.790, 95% CI [2.697, 2.885], and 1.417, 95% CI [1.359, 1.477], respectively. For the comparison group, the estimated odds ratios for large compared to no More at Four expenditures, for moderate compared to no More at Four expenditures, and for small compared to no More at Four expenditures were 5.021, 95% CI [4.897, 5.148], 2.415, 95% CI [2.365, 2.465], and 1.291, 95% CI [1.259, 1.325], respectively.

The positive effects of More at Four expenditures on consistently competent functioning overall as well as within the academic performance and behavioral functioning domains across time were not found with respect to the special education domain. There was a slight negative effect of large compared to no expenditures (*IOR* = 1.124, 95% CI [1.099, 1.151]) and moderate compared to no expenditures (*IOR* = 1.106, 95% CI [1.085, 1.129]). However, this effect was small, $F(3, 302) = 53.530$, $p < .0001$, indicating that there was likely not a meaningful difference across levels of More at Four expenditures on competence within the special education domain. The interaction between county expenditures on More at Four and maltreatment status was not significant, $F(3, 299) = 2.320$, $p = .076$, indicating no meaningful differences across groups.

In multilevel regression analyses including More at Four expenditures as a Level 1 predictor as well as the covariates, living in a county with relatively large expenditures on More at Four predicted being competent across domains for a greater percentage of years, $F(3, 302) = 12{,}035.600$, $p < .0001$. The effect of maltreatment status on the percentage of years in which participants were competent was significant after controlling for the effects of More at Four as well as the covariates, $F(1, 101) = 11{,}894.700$, $p < .0001$. In addition, the interaction between More at Four expenditures and maltreatment status was significant but small, $F(3, 299) = 23.780$, $p < .0001$, indicating that there might be differences in the effects of More at Four expenditures across groups. The small size of the effect suggested that group differences were not likely to be substantial.

Examination of the model parameter estimates indicated that maltreated participants in counties with large, moderate, and small levels of More at Four expenditures were slightly more likely to be competent over a larger percentage of years than nonmaltreated participants in counties with similar levels of expenditures, suggesting a modest positive differential impact on maltreated participants from More at Four expenditures. In analyses by subgroup, relative levels of expenditures on More at Four predicted the percentage of years in which participants were competent for both the maltreated group, $F(3, 300) = 3{,}709.160$, $p < .0001$, and the comparison group, $F(3, 302) = 12{,}035.600$, $p < .0001$.

3.3.8.3 Separate effects of Smart Start on consistent competence

Supplemental analyses of county expenditures were conducted on the Smart Start program excluding More at Four expenditures. Total per capita expenditures on Smart Start were categorized as small (between $500 and $1,000, approximately 9% of each group), moderate (between $1,000 and $1,500, approximately 36% of the maltreated group and 39% of the comparison group), moderately large (between $1,500 and $2,000, approximately 29% of each

group), and large (greater than $2,000, approximately 27% of the maltreated group and 22% of the comparison group). These relative percentages were consistent across the three different subsamples of movers and nonmovers within each group used in these analyses. Prior to conducting multilevel analyses, the relative percentages of competent and non-competent participants at each level of county expenditures on Smart Start were calculated. Among participants living in counties with small expenditures on Smart Start, 23% were competent, whereas among participants living in counties with large expenditures on Smart Start and More at Four, 26% were competent. Thirty-three percent (33%) of participants living in counties with moderately large expenditures on Smart Start were competent and 35% of participants living in counties with moderate expenditures on Smart Start were competent. Figure 5 depicts the interaction of Smart Start expenditures by maltreatment status on participants' consistent competence without adjusting for any covariates. As shown in Figure 5, the relation between relative expenditures on Smart Start and consistent competence is curvilinear rather linear, with a pattern of increasing competence across small to moderate levels of expenditures, relative flattening of competence at moderately large levels of expenditures, and apparent diminishing returns at the highest level of expenditures on Smart Start. This pattern of effects was found to be robust across subsequent analyses controlling for covariates. Relative levels of government expenditures on Smart Start predicted consistent competence overall ($\chi^2(3) = 5{,}654.677$, $p < .0001$).

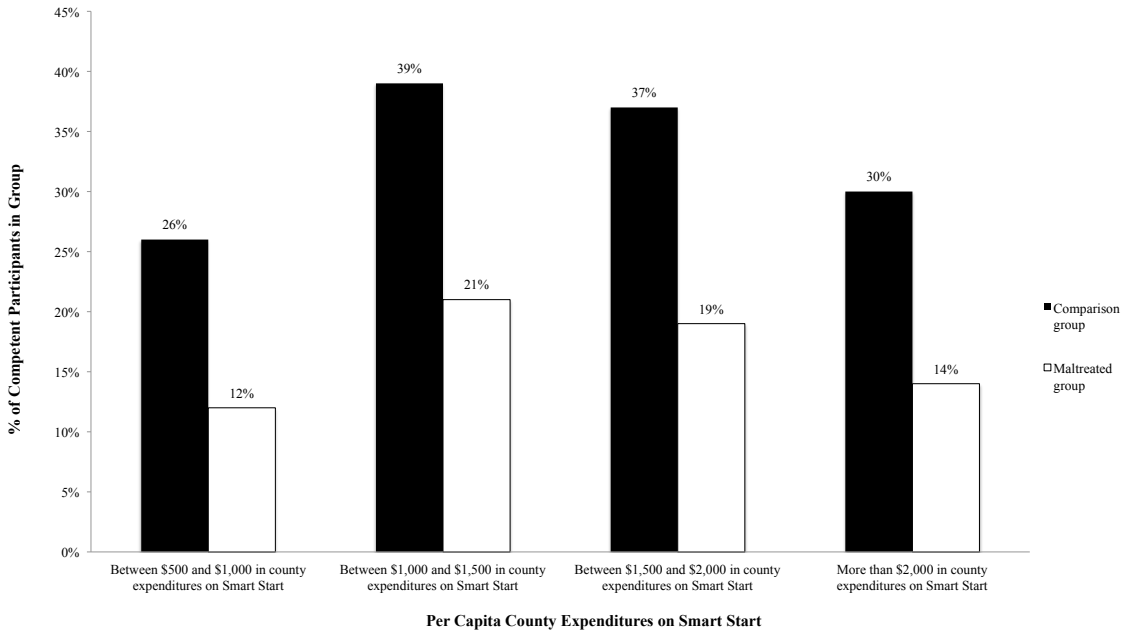

Figure 5: Consistent Competence by Expenditures on Smart Start and Maltreatment Status

The overall effect of government expenditures on Smart Start on participants' competence was examined using multilevel logistic regression. The first set of analyses included all participants assuming movers remained in their birth county during the ages of 0-5 and thus received the benefit of expenditures on Smart Start in their birth county. Relative levels of Smart Start expenditures predicted consistent competence, $F(3, 303) = 1{,}911.750$, $p < .0001$. After adjusting for covariates as well as the effect of maltreatment status and the interaction of Smart Start expenditures and maltreatment status, participants born in counties with large expenditures on Smart Start had significantly greater odds of being competent than participants born in counties with small expenditures on Smart Start (*OR* = 1.381, 95% CI [1.328, 1.435]). Participants who were born in counties with moderate expenditures or moderately large expenditures on Smart Start during the years in which participants were 0-5 years of age had more than twice the odds of being consistently competent as participants who were born in counties with relatively small expenditures on Smart Start (*OR* = 2.241, 95% CI [2.161, 2.324],

for moderately large versus small expenditures; *OR* = 2.347, 95% CI [2.269, 2.428], for moderate versus small expenditures). Participants in the comparison group had 1.5 times the odds of being consistently competent as participants in the maltreated group (*OR* = 1.470, 95% CI [1.440, 1.501]). The interaction of Smart Start expenditures and maltreatment status was significant but small, $F(3, 298) = 5.73$, $p = .001$, indicating that there might be differences in the effect of Smart Start expenditures on consistent competence among participants in the maltreated and comparison groups.

Alternative analyses examined the effect of Smart Start expenditures using the full sample assuming that movers lived in their school county from age 0-5. The results were substantially similar to those for the full sample assuming movers remained in their county of birth from age 0-5. Another set of analyses used only the nonmover sample of participants who were assumed to remain in their birth county between their date of birth and the date they entered school. The pattern of results was virtually identical to the results reported above for the full sample including movers.

As the foregoing analyses using the full sample indicated that the effects of Smart Start expenditures might be different across groups, subsequent analyses examined the effects of expenditures on Smart Start within the maltreated group and comparison group separately. Smart Start expenditures were found to be a significant predictor of consistent competence among maltreated participants, $F(3, 299) = 510.310$, $p < .0001$. After adjusting for covariates, maltreated participants born in counties with large expenditures on Smart Start had significantly greater odds of being competent than maltreated participants born in counties with small expenditures on Smart Start (*OR* = 1.204, 95% CI [1.120, 1.294]). Participants in the maltreated group who were born in counties with moderate expenditures or moderately large expenditures on Smart Start during the years in which participants were 0-5 years of age had twice the odds of being

consistently competent as participants in the maltreated group who were born in counties with relatively small expenditures on Smart Start (*OR* = 2.025, 95% CI [1.892, 2.167], for moderately large versus small expenditures; *OR* = 2.190, 95% CI [2.059, 2.330], for moderate versus small expenditures).

Alternative analyses examined the effect of Smart Start expenditures using the full maltreated group assuming that movers lived in their school county from age 0-5. The results were substantially similar to those for the full maltreated sample assuming movers remained in their county of birth from age 0-5, except that maltreated participants whose county made large expenditures on Smart Start had equivalent odds of being competent as participants in the maltreated group living in counties with small expenditures on Smart Start (*OR* = 1.055, 95% CI [.967, 1.150]), whereas participants in the maltreated group living in counties with moderate or moderately large expenditures on Smart Start had more than two times the odds of being competent as nonmovers in the maltreated group living in counties with small expenditures on Smart Start (*OR* = 2.081, 95% CI [1.928, 2.247]), for moderately large versus small expenditures; *OR* = 2.353, 95% CI [2.208, 2.508], for moderate versus small expenditures)

Another set of analyses used only the nonmover sample of maltreated participants who were assumed to remain in their birth county between their date of birth and the date they entered school, with similar results to those reported with respect to analyses using the full maltreated sample. Similar to the results in the primary analyses, nonmovers in the maltreated group living in counties with moderately large expenditures on Smart Start had almost twice the odds of being competent as nonmovers in the maltreated group living in counties with small expenditures on Smart Start (*OR* = 1.876, 95% CI [1.700, 2.070]), and nonmovers in the maltreated group living in counties with moderate expenditures on Smart Start had almost 2.5 times the odds of being competent as nonmovers in the maltreated group living in counties with small expenditures on

Smart Start (*OR* = 2.403, 95% CI [2.222, 2.599]). However, participants living in counties with small expenditures on Smart Start had greater odds of being consistently competent than participants living in counties with large expenditures on Smart Start (*IOR* = 1.156, 95% CI [1.031, 1.295]).

The effect of Smart Start expenditures on consistent competence also was examined for participants in the comparison group to determine whether there was any differential impact on maltreated participants. The pattern of results was similar and perhaps slightly larger for participants in the comparison group. Smart Start expenditures predicted consistent competence among nonmaltreated participants, $F(3, 302) = 1{,}372.720$, $p < .0001$. After adjusting for covariates, nonmaltreated participants born in counties with large expenditures on Smart Start had 1.5 times the odds of being competent than nonmaltreated participants born in counties with small expenditures on Smart Start (*OR* = 1.487, 95% CI [1.436, 1.541]). Participants in the comparison group who were born in counties with moderate expenditures or moderately large expenditures on Smart Start during the years in which participants were 0-5 years of age had almost 2.5 times the odds of being consistently competent as participants in the comparison group who were born in counties with relatively small expenditures on Smart Start (*OR* = 2.314, 95% CI [2.240, 2.390], for moderately large versus small expenditures; *OR* = 2.424, 95% CI [2.354, 2.495], for moderate versus small expenditures). The pattern of results for the effects of Smart Start expenditures on nonmaltreated participants' consistently competent functioning was similar across analyses including all participants in the comparison group assuming that the movers moved to their school county after birth as well as the nonmover subsample of the comparison group.

Table 23 describes the odds ratios and 95% confidence intervals for the fixed effects in three fitted models for the effects of Smart Start expenditures for maltreated participants using the full maltreated group assuming movers remained in their birth county in the relevant years, the

full maltreated group assuming movers moved to their school county immediately after birth, and the subsample of nonmovers in the maltreated group, respectively. Table 24 describes the odds ratios and 95% confidence intervals for the fixed effects in three fitted models for the effects of Smart Start expenditures for the full sample of participants in the comparison group assuming movers remained in their birth county in the relevant years, the full sample of participants in the comparison group assuming movers moved to their school county immediately after their birth, and the subsample of nonmovers in the comparison group, respectively.

Table 23: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start - Maltreated Group)

Variable	Include movers in birth county		Include movers in school county		Nonmovers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
Smart Start county expenditures						
Investment > $2,000 vs. $500 - $1,000	1.081	[1.008, 1.159]	1.164 *(IOR)*	[1.072, 1.264]	1.550 *(IOR)*	[1.391, 1.730]
Investment $1,500 - $2,000 vs. $500 - $1,000	1.974	[1.849, 2.108]	1.915	[1.779, 2.062]	1.635	[1.487, 1.798]
Investment $1,000 - $1,500 vs. $500 - $1,000	2.209	[2.081, 2.344]	2.321	[2.183, 2.467]	2.359	[2.188, 2.544]
			Model 3			
Smart Start county expenditures						
Investment > $2,000 vs. $500 - $1,000	1.204	[1.120, 1.294]	1.055	[.967, 1.150]	1.156 *(IOR)*	[1.031, 1.295]
Investment $1,500 - $2,000 vs. $500 - $1,000	2.025	[1.892, 2.167]	2.081	[1.928, 2.247]	1.876	[1.700, 2.070]
Investment $1,000 - $1,500 vs. $500 - $1,000	2.190	[2.059, 2.330]	2.353	[2.208, 2.508]	2.403	[2.222, 2.599]
Gender						
Female vs. male	2.020	[1.963, 2.079]	2.021	[1.964, 2.080]	1.992	[1.926, 2.060]
Race/ethnicity						
White vs. Black	2.237 *(IOR)*	[2.151, 2.326]	2.242 *(IOR)*	[2.155, 2.331]	2.304 *(IOR)*	[2.203, 2.415]
White vs. Hispanic	1.736 *(IOR)*	[1.608, 1.873]	1.733 *(IOR)*	[1.608, 1.869]	1.783 *(IOR)*	[1.634, 1.946]
White vs. Other	1.183 *(IOR)*	[1.130, 1.241]	1.189 *(IOR)*	[1.134, 1.245]	1.202 *(IOR)*	[1.136, 1.272]
SES						
High SES vs. low SES	3.729	[3.557, 3.910]	3.719	[3.547, 3.900]	3.801	[3.593, 4.021]
Middle SES vs. low SES	2.146	[2.077, 2.219]	2.144	[2.127, 2.203]	2.206	[2.121, 2.295]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.122	[1.078, 1.169]	1.122	[1.077, 1.168]	1.117	[1.065, 1.172]
Mother's marital status						
Married vs. not married	1.059	[1.026, 1.094]	1.063	[1.029, 1.097]	1.058	[1.018, 1.099]
Birth year (GMC, RC) X years of education data (GMC)	1.056	[1.052, 1.059]	1.053	[1.050, 1.057]	1.055	[1.051, 1.060]
Mother's age at birth (GMC)	1.007 *(IOR)*	[1.004, 1.009]	1.006 *(IOR)*	[1.004, 1.009]	1.006 *(IOR)*	[1.003, 1.009]

Note. Samples used in analyses were: full maltreated sample assuming movers lived in their birth county during ages 0-5, full maltreated sample assuming movers lived in the county in which they attended school for most or all of the years of available education data during ages 0-5, and subsample of maltreated participants who remained in same county at birth and during most or all education years (nonmovers) and were assumed to have lived in their birth county during ages 0-5. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start within the county during years participants were age 0-5. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

Table 24: Odds Ratios and 95% Confidence Intervals for Multilevel Models Predicting Consistent Competence for Children Nested Within School County (Smart Start - Comparison Group)

Variable	Include movers in birth county		Include movers in school county		Nonmovers	
	OR/IOR	95% CI	*OR/IOR*	95% CI	*OR/IOR*	95% CI
			Model 2			
Smart Start county expenditures						
Investment > $2,000 vs. $500 - $1,000	1.290	[1.249, 1.333]	1.059	[1.015, 1.104]	1.012 *(IOR)*	[.962, 1.066]
Investment $1,500 - $2,000 vs. $500 - $1,000	1.969	[1.911, 2.029]	1.917	[1.851, 1.986]	1.855	[1.778, 1.936]
Investment $1,000 - $1,500 vs. $500 - $1,000	2.040	[1.985, 2.096]	2.116	[2.059, 2.175]	2.120	[2.051, 2.191]
			Model 3			
Smart Start county expenditures						
Investment > $2,000 vs. $500 - $1,000	1.487	[1.436, 1.541]	1.279	[1.221, 1.339]	1.200	[1.133, 1.270]
Investment $1,500 - $2,000 vs. $500 - $1,000	2.314	[2.240, 2.390]	2.439	[2.347, 2.531]	2.399	[2.289, 2.514]
Investment $1,000 - $1,500 vs. $500 - $1,000	2.424	[2.354, 2.495]	2.607	[2.531, 2.686]	2.660	[2.567, 2.757]
Gender						
Female vs. male	1.923	[1.896, 1.949]	1.925	[1.899, 1.952]	1.915	[1.884, 1.946]
Race/ethnicity						
White vs. Black	2.331 *(IOR)*	[2.283, 2.381]	2.342 *(IOR)*	[2.294, 2.392]	2.415 *(IOR)*	[2.358, 2.481]
White vs. Hispanic	2.183 *(IOR)*	[2.110, 2.257]	2.203 *(IOR)*	[2.132, 2.278]	2.273 *(IOR)*	[2.188, 2.358]
White vs. Other	1.149 *(IOR)*	[1.120, 1.181]	1.161 *(IOR)*	[1.131, 1.192]	1.176 *(IOR)*	[1.142, 1.214]
SES						
High SES vs. low SES	3.924	[3.841, 4.009]	3.940	[3.856, 4.026]	3.922	[3.824, 4.021]
Middle SES vs. low SES	2.038	[1.996, 2.081]	2.041	[1.998, 2.084]	2.030	[1.981, 2.080]
Fetal well-being						
Full-term and normal BW vs. premature and low BW	1.143	[1.118, 1.168]	1.147	[1.122, 1.172]	1.142	[1.114, 1.171]
Mother's marital status						
Married vs. not married	1.141	[1.119, 1.164]	1.144	[1.121, 1.167]	1.134	[1.107, 1.161]
Birth year (GMC, RC) X years of education data (GMC)	1.049	[1.047, 1.052]	1.049	[1.046, 1.051]	1.047	[1.045, 1.050]
Mother's age at birth (GMC)	1.009	[1.008, 1.010]	1.009	[1.008, 1.010]	1.010	[1.008, 1.012]

Note. Samples used in analyses were: full comparison group assuming movers lived in their birth county during ages 0-5, full full comparison group assuming movers lived in the county in which they attended school for most or all of the years of available education data during ages 0-5, and subsample of full comparison group who remained in same county at birth and during most or all education years (nonmovers) and were assumed to have lived in their birth county during ages 0-5. Model 1 (not shown) is an intercept-only model with one Level 2 variable, county, and no Level 1 predictors. Model 2 includes a Level 1 predictor that reflects per capita county expenditures on Smart Start within the county during years participants were age 0-5. Model 3 includes several Level 1 covariates, such as gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, birth year, and years of education data. BW = birth weight; SES = socioeconomic status; GMC = grand mean centered; RC = reverse coded; CI = confidence interval; OR = odds ratio; IOR = inverse odds ratio. IOR is equal to (1 / odds ratio), calculated for odds ratios less than 1 for ease of interpretation. Interpretation of IOR is inverse of odds ratio (e.g., reference group has [IOR] times the odds of being competent as comparison group).

3.3.8.4 Separate effects of Smart Start on patterns of competence

In addition to the supplemental multilevel analyses investigating the separate effects of county expenditures on Smart Start on participants' consistently competent functioning across time, additional multilevel analyses examined the effects of the relative levels of expenditures on Smart Start on patterns of competence among participants. Multilevel logistic regression models were used to investigate effects on consistent competence across time within specific domains of functioning, including academic performance, special education, and behavioral functioning domains. Multilevel regression analyses were used to examine effects on the percentage of years in which participants demonstrated competence across all domains.

After controlling for gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of education data as well as maltreatment status, relative expenditures on Smart Start within participants' county when they were age 0-5 predicted competence across time in the domain of academic functioning, $F(3, 303) = 2{,}165.570$, $p < .0001$. Participants who lived in counties with relatively large expenditures on Smart Start had more than 1.25 times the odds of being competent in the academic performance domain than participants who lived in counties with small expenditures on Smart Start (*OR* = 1.272, 95% CI [1.233, 1.313]), and participants who lived in counties with moderately large levels of expenditures had more than twice the odds of being competent in the academic performance domain than participants who lived in counties with small expenditures (*OR* = 2.105, 95% CI [2.043, 2.168]). Participants who lived in counties with moderate levels of expenditures on Smart Start had almost 2.5 times the odds of being competent in the academic performance domain than participants who lived in counties with small Smart Start expenditures (*OR* = 2.355, 95% CI [2.291, 2.419]).

The effect of maltreatment status on academic performance remained significant after controlling for the effect of Smart Start as well as the covariates, $F(1, 101) = 524.540$, $p < .0001$, such that nonmaltreated participants had somewhat greater odds of functioning competently in the academic performance domain across time than maltreated participants ($OR = 1.216$, 95% CI [1.196, 1.237]). In addition, the interaction between Smart Start expenditures and maltreatment status was significant but small, $F(3, 298) = 23.460$, $p < .0001$, indicating that there might be differences in the effects of expenditures on Smart Start on competence within the academic performance domain across groups. The small size of the effect suggested that group differences were not likely to be substantial.

In analyses run separately by group, the effects of relative expenditures on Smart Start on competence within the academic performance domain were similar across groups but somewhat larger in the comparison group than in the maltreated group. For the comparison group, the estimated odds ratios for large compared to small Smart Start expenditures, for moderately large compared to small Smart Start expenditures, and for moderate compared to small Smart Start expenditures were 1.451, 95% CI [1.406, 1.499], 2.374, 95% CI [2.304, 2.446], and 2.606, 95% CI [2.537, 2.677], respectively. For the maltreated group, the estimated odds ratios for large compared to small expenditures, for moderately large compared to small expenditures, and for moderate compared to small expenditures were 1.089, 95% CI [1.030, 1.151], 1.825, 95% CI [1.732, 1.923], and 2.106, 95% CI [2.008, 2.208], respectively.

Results for analyses examining effects of Smart Start expenditures on consistently competent functioning within the behavioral domain were similar to those reported for the academic performance domain. The effect of maltreatment status on behavioral functioning was significant after controlling for the effects of county expenditures and the covariates, $F(1, 101) = 4{,}444.800$, $p < .0001$, such that nonmaltreated participants had greater odds of functioning

competently in the behavioral domain across time than maltreated participants (*OR* = 1.751, 95% CI [1.722, 1.781]). The interaction between Smart Start expenditures and maltreatment status was significant, $F(3, 298) = 13.580$, $p < .0001$, indicating potential differences across groups in the effects of Smart Start expenditures on competence within the behavioral functioning domain. The small size of the effect suggested that any group differences were likely not substantial.

In analyses run separately by group, the effects of relative levels of Smart Start expenditures on competence within the behavioral functioning domain were similar across groups. For the maltreated group, the estimated odds ratios for large compared to small expenditures, for moderately large compared to small expenditures, and for moderate compared to small expenditures were 1.398, 95% CI [1.326, 1.475], 2.296, 95% CI [2.183, 2.414], and 2.553, 95% CI [2.442, 2.670], respectively. For the comparison group, the estimated odds ratios for large compared to small Smart Start expenditures, for moderately large compared to small Smart Start expenditures, and for moderate compared to small Smart Start expenditures were 1.310, 95% CI [1.267, 1.355], 2.124, 95% CI [2.058, 2.191], and 2.411, 95% CI [2.345, 2.478], respectively.

The positive effects of Smart Start expenditures on consistently competent functioning overall as well as within the academic performance and behavioral functioning domains across time were not found with respect to the special education domain. There was a slight negative effect of large compared to small expenditures (*IOR* = 1.045, 95% CI [1.012, 1.080]), moderately large compared to small expenditures (*IOR* = 1.115, 95% CI [1.081, 1.149]), and moderate compared to small expenditures (*IOR* = 1.086, 95% CI [1.057, 1.117]). The effect of maltreatment status on competence in the special education domain was significant after controlling for the effects of Smart Start expenditures and the covariates, $F(1, 101) = 846.050$, $p < .0001$, such that nonmaltreated participants had greater odds of functioning competently in the special education domain across time than maltreated participants (*OR* = 1.293, 95% CI [1.271,

1.316]). The interaction between county expenditures on Smart Start and maltreatment status was not significant, $F(3, 298) = 2.190$, $p = .090$, indicating no meaningful differences in the effects of Smart Start expenditures existed across groups.

In multilevel regression analyses including Smart Start expenditures as a Level 1 predictor as well as the covariates, living in a county with relatively large expenditures on Smart Start predicted being competent across domains for a larger percentage of years, $F(3, 303) = 3{,}196.340$, $p < .0001$. The effect of maltreatment status on the percentage of years in which participants were competent was significant after controlling for the effects of Smart Start as well as the covariates, $F(1, 101) = 10{,}473.000$, $p < .0001$. In addition, the interaction between Smart Start expenditures and maltreatment status was significant but small, $F(3, 298) = 4.380$, $p = .005$, indicating that there might be differences in the effects of Smart Start expenditures across groups. The small size of the effect suggested that group differences were not likely to be substantial. In analyses by subgroup, relative levels of expenditures on Smart Start predicted the percentage of years in which participants were competent for both the maltreated group, $F(3, 299) = 945.270$, $p < .0001$, and the comparison group, $F(3, 302) = 2{,}925.090$, $p < .0001$.

3.4 Multilevel Analyses of the Independent Effects of Community-Level Factors

Supplemental analyses examined whether there were independent effects of the two community-level factors investigated in the current study, specifically, the introduction of the Multiple Response System within counties and government expenditures on Smart Start and More at Four. Prior to conducting multilevel analyses, the relative percentages of consistently competent and non-competent participants by level of county expenditures and date of introduction of the Multiple Response System in their county were calculated. As shown in Table 25, rates of competence were higher among participants whose county had adopted the Multiple

Response System prior to the date their maltreatment report was investigated. Rates of competent also were higher at larger relative levels of expenditures on Smart Start and More at Four. A total of 29% of participants who lived in counties with relatively large expenditures on Smart Start and More at Four and were reported for maltreatment after the Multiple Response System had been adopted in their county were consistently competent, compared to 16% of participants who were reported for maltreatment prior to the introduction of the Multiple Response System in their county. A total of 23% of participants who lived in counties with moderate expenditures on Smart Start and More at Four and were reported for maltreatment after the introduction of the Multiple Response System in their county were consistently competent, compared to 12% of participants reported for maltreatment prior to the introduction of the Multiple Response System in their county. Similarly, 21% of participants who lived in counties with small expenditures on Smart Start and More at Four and were reported for maltreatment after the introduction of the Multiple Response System in their county were consistently competent, compared to only 10% of participants reported for maltreatment prior to their county's adoption of the Multiple Response System.

Table 25: Consistent Competence by Government Expenditures on Smart Start and More at Four and Date of Introduction of Multiple Response System

Variable	Maltreatment Investigated Prior to Introduction of Multiple Response System	Maltreatment Investigated After Introduction of Multiple Response System
Smart Start and More at Four county expenditures		
Investment > $2,500	16% competent	29% competent
Investment $1,500 - $2,500	12% competent	23% competent
Investment $500 - $1,500	10% competent	21% competent

Multilevel logistic regression models in which both factors were entered simultaneously produced similar results to analyses in which each factor was analyzed separately, indicating independent and additive effects of each program. For example, in a model predicting consistent competence among the full maltreated sample assuming movers remained in their birth county prior to school entry and adding both factors as Level 1 predictors to the unconditional model, the estimated odds ratio associated with the introduction of the Multiple Response System was 2.290, 95% CI [2.226, 2.357] (compared to *OR* = 2.402, 95% CI [2.335, 2.471], in a single-predictor model). Similarly, the estimated odds ratio for large compared to small expenditures on early childhood programs was 2.244, 95% CI [2.133, 2.360] (compared to *OR* = 2.593, 95% CI [2.466, 2.725], in a single-predictor model), and the estimated odds ratio for moderate compared to small investments in early childhood programs was 1.613, 95% CI [1.549, 1.679] (compared to *OR* = 1.810, 95% CI [1.739, 1.883], in a single-predictor model).

Each Level 1 factor continued to predict consistent competence after controlling for the effects of covariates (e.g., *OR* = 2.026, 95% CI [1.968, 2.086], for maltreated participants reported for maltreatment after introduction of the Multiple Response System in their county compared to participants reported prior to its introduction; *OR* = 2.090, 95% CI [1.982, 2.204], for large compared to small expenditures on Smart Start and More at Four; *OR* = 1.535, 95% CI [1.472, 1.602], for moderate compared to small expenditures on Smart Start and More at Four). Results were similar to those presented in the previous analyses focused on each individual community-level predictor with respect to each sample of movers and nonmovers and across maltreated and comparison groups for combined government expenditures on Smart Start and More at Four as well as on each early childhood program analyzed separately.

3.5 Analyses of Patterns of Competent Functioning Before and After Experiences of Maltreatment

Additional analyses examined the competent functioning of maltreated children in years prior to experiences of maltreatment. A subsample of maltreated participants had years of education data available prior to the academic year in which they were reported for maltreatment. In the primary analyses of consistent competence following maltreatment, September 1 was used as a reference date for the beginning of each academic year, and the academic year falling subsequent to the date participants were reported for maltreatment was deemed to be the first eligible year of data. This conservative criterion was employed to limit outcomes to those occurring after maltreatment occurred and was investigated. For similar reasons, in the current set of analyses on competent functioning prior to maltreatment, the academic year in which participants were reported for maltreatment was excluded in order to limit outcomes to those reasonably assumed to have occurred prior to experiences of maltreatment, though it is possible that participants may have been maltreated long before a report was made and investigated by CPS.

As previously reported above, a majority of the maltreated group (n = 96,184) were reported for maltreatment prior to the beginning of the first year of education data used in the analyses, and all of their education data were considered to have been measured post-maltreatment. Of the remaining 57,313 maltreated participants, a total of 45,986 participants had years of available education data prior to their first eligible year of data for the primary analyses, and a total of 34,131 participants had years of available education data prior to the year in which they were reported for maltreatment. This subsample of 34,131 was used in the analyses of pre-maltreatment competence.

As a result of the fact that this subsample was limited to those children who had already started public school prior to being reported for maltreatment, participants in the subsample were

older when they were first reported for maltreatment than participants in the full maltreated sample, with an average age of 9 years old at the time of their maltreatment report (M = 8.892, SD = 2.173) compared to 4 years old in the full maltreated sample. The average number of available years of data prior to the year in which they were reported for maltreatment was 2.462 (SD = 1.309), ranging from 1 to 5 years, and the average number of years of available data following experiences of maltreatment was 2.806 (SD = 1.362), ranging from 1 to 5 years. The average number of total available years of pre-maltreatment and post-maltreatment data for the subsample (excluding the academic year in which participants were reported for maltreatment) was 5.268 (SD = 1.044), with a minimum of 2 and maximum of 6 years out of the total 7 years included in the current study.

3.5.1 Prevalence of Consistent Competence During Years Prior to and Following Experiences of Maltreatment

Using the same competence criteria as in the primary analyses, a total of 55% of the subsample were found to be consistently competent within all domains across all years of pre-maltreatment education data. A total of 61% were consistently competent within the academic performance domain, 85% were consistently competent within the special education domain, and 85% were consistently competent within the behavioral functioning domain across pre-maltreatment years. In examining the total number of domains in which participants were competent over time during the years prior to being reported for maltreatment, 46% were consistently competent in three domains, 36% were consistently competent in two domains, 15% were consistently competent in one domain, and only 3% were never consistently competent in any domain prior to being reported for maltreatment.

Additional analyses examined the subsample's functioning post-maltreatment. A total of 25% were consistently competent within all domains across years of data following the

experience of maltreatment, compared to 18% in the full maltreated sample. A total of 45% were consistently competent within the academic performance domain, 72% were consistently competent within the special education domain, and 54% were consistently competent within the behavioral functioning domain across post-maltreatment years. Of those who were functioning competently in the academic performance domain prior to being reported for maltreatment, 53% continued to be consistently competent within that domain after their maltreatment experience. Of those who were functioning competently in the special education domain prior to experiencing maltreatment, 82% continued to be consistently competent within that domain following maltreatment. Of those who were functioning competently in the behavioral domain prior to experiencing maltreatment, 58% continued to be consistently competent within that domain after being investigated for maltreatment.

Participants' overall pattern of functioning across pre-maltreatment and post-maltreatment years was assessed. Among participants who were consistently competent before they were reported for maltreatment, only about a third continued to be consistently competent following the experience of maltreatment, resulting in 19% of the subsample being deemed "always competent" across years of eligible data. This subgroup reflects a consistent pattern of competent functioning across all time points measured for the current study, excluding the year in which participants were reported for maltreatment. The remaining two-thirds of participants who were consistently competent before they were reported for maltreatment were not consistently competent following maltreatment and were deemed to have "declined" in competent functioning across years of eligible data. This subgroup, comprising 36% of the subsample, reflects a discontinuity in competence following the experience of maltreatment.

Among participants who were not consistently competent in the years before they were reported for maltreatment, 85% continued not to be competent after maltreatment, resulting in

38% of the subsample being deemed to be "never competent" across years of eligible data. This subgroup represents the participants demonstrating the lowest functioning across time within the maltreated group. Among participants who were not consistently competent in the years before they were reported for maltreatment, only 15% demonstrated consistent competence following the experience of maltreatment, resulting in 7% of the subsample being deemed to have "improved" in competent functioning across years of eligible data. This subgroup comprises the relatively small group of participants who demonstrated competent functioning following the experience of maltreatment despite exhibiting relatively poor functioning prior to being reported for maltreatment.

Though data from the year in which participants were reported for maltreatment were not included in the competence analyses for the reasons described above, it was determined that 88% of participants in the "always competent" category also were competent across domains in the maltreatment year as well as 61% of participants in the "improved" subgroup and 64% of participants in the "declined" category. Twenty-one (21%) of participants in the "never competent" subgroup were competent across domains during the year in which they were reported for maltreatment. Data on competence during the maltreatment year were missing for 142 participants in these analyses. Figure 6 shows the relative percentages of maltreated participants in each category of competent functioning prior to and following experiences of maltreatment (never competent, declined in competence, always competent, and improved in competence). Results are presented for the overall subsample used in these analyses as well as by age at which participants were first reported for maltreatment.

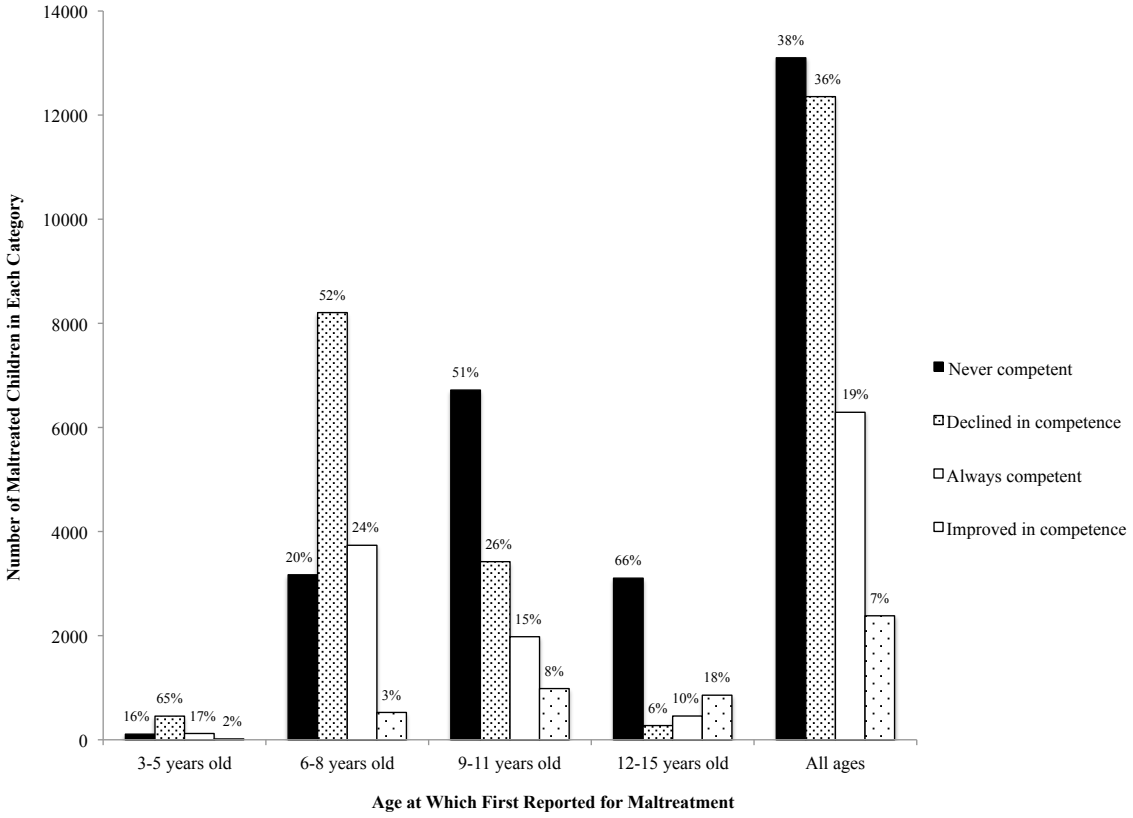

Figure 6: Competence Across Pre-Maltreatment and Post-Maltreatment Years by Age at Maltreatment

3.5.2 Effects of Government Expenditures on Smart Start and More at Four on Competence Prior to and Following Experiences of Maltreatment

In addition to examining the prevalence of consistent competence for maltreated participants in the years prior to being reported for maltreatment, supplemental analyses investigated whether county expenditures on the Smart Start and More at Four programs predicted participants' patterns of competence. Three sets of analyses were conducted with respect to investigating potential program effects on competence. The first set of analyses tested whether relative levels of expenditures on Smart Start and More at Four predicted participants' pre-maltreatment competence. The second set of analyses investigated whether relative levels of expenditures on Smart Start and More at Four predicted post-maltreatment competence for this subgroup of participants. The third set of analyses examined whether relative levels of

expenditures on Smart Start and More at Four predicted participants' consistent competence across both pre-maltreatment and post-maltreatment time periods, which was operationalized as being classified in the "always competent" subgroup.

Multilevel logistic regression analyses were conducted using the subsample of maltreated participants with available education data prior to the year in which they were reported for maltreatment to investigate whether expenditures on early childhood programs during the years that children were age 0-5 predicted competent outcomes in the relevant time periods. The analyses were conducted using the same basic models as those used in the previous analyses of the effects of Smart Start and More at Four, except that in analyses of competence in the years prior to being reported for maltreatment, the relevant covariate used to represent participants' years of available data was the years of pre-maltreatment data. The independent and joint effects of the introduction of the Multiple Response System could not be evaluated in these analyses, as all of the maltreated participants in the subsample were reported for maltreatment after the Multiple Response System had been adopted in the county investigating their maltreatment report.

First, the overall effect of county expenditures on Smart Start and More at Four on participants' consistent competence within all domains and across all pre-maltreatment years was analyzed. After controlling for gender, race/ethnicity, socioeconomic status, fetal well-being, mother's marital status at birth, mother's age at birth, participant's birth year, and years of pre-maltreatment education data, relative expenditures on Smart Start and More at Four within participants' county when they were age 0-5 predicted consistent competence prior to the year in which participants were reported for maltreatment, $F(2, 196) = 310.250$, $p < .0001$. Participants who lived in counties with relatively large expenditures on Smart Start and More at Four had three times the odds of being competent across pre-maltreatment years than participants who lived in counties with small expenditures on Smart Start and More at Four ($OR = 2.995$, 95% CI

[2.728, 3.289]), and participants who lived in counties with moderate levels of expenditures had more than twice the odds of being competent across pre-maltreatment years than participants who lived in counties with small expenditures (*OR* = 2.209, 95% CI [2.058, 2.374]).

In analyses examining the effects of county expenditures for each program separately, similar results were found with respect to both More at Four and Smart Start. After controlling for the effects of the covariates, relative expenditures on More at Four within participants' county when they were age 4 predicted consistent competence across pre-maltreatment years, $F(3, 290) = 214.660$, $p < .0001$. Participants who lived in counties with large or moderate More at Four expenditures had almost four times the odds of being competent prior to experiencing maltreatment than participants who lived in counties with no expenditures on More at Four (*OR* = 3.765, 95% CI [3.277, 4.325], and *OR* = 3.967, 95% CI [3.561, 4.419], respectively), and participants who lived in counties with small levels of More at Four expenditures had more than twice the odds of being competent in the pre-maltreatment period than participants who lived in counties with no More at Four expenditures (*OR* = 2.266, 95% CI [2.052, 2.503]).

Similarly, relative expenditures on Smart Start within participants' county when they were age 0-5 predicted consistent competence across pre-maltreatment years after controlling for the effects of covariates, $F(3, 290) = 220.320$, $p < .0001$. Participants who lived in counties with large expenditures on Smart Start in the years in which participants were age 0-5 had three times the odds of being competent prior to experiencing maltreatment than participants who lived in counties with small expenditures on Smart Start (*OR* = 2.954, 95% CI [2.615, 3.336]). Participants who lived in counties with relatively large or moderate levels of Smart Start expenditures had four times the odds of being competent across pre-maltreatment years than participants who lived in counties with small levels of Smart Start expenditures (*OR* = 4.171, 95% CI [3.700, 4.702], and *OR* = 3.995, 95% CI [3.567, 4.474], respectively).

Second, the effects of expenditures on Smart Start and More at Four on subsample participants' post-maltreatment competence were investigated. Post-maltreatment competence was the same outcome as these participants' consistent competence in the primary analyses, where only years following the year in which participants were reported for maltreatment were included. Thus, this set of analyses effectively examined the effects of Smart Start and More at Four expenditures on the identified subsample of maltreated participants. Though the model including only Smart Start and More at Four expenditures as a Level 1 predictor and participants' school county as a Level 2 predictor of consistent competence indicated a significant effect, in the model adjusted for covariates, relative expenditures on Smart Start and More at Four within participants' county when they were age 0-5 did not meaningfully predict consistent competence following maltreatment, $F(2, 196) = 4.200$, $p = .016$. Though the overall fixed effect was significant, the size of the effect was quite small, and the confidence intervals around the odds ratios for each level of Smart Start and More at Four expenditures included 1.000, indicating that the effects on competence were equivalent across large compared to small and moderate compared to small government expenditures. Examination of the model parameters suggests that the effect of large compared to moderate expenditures may be significant and positive, though likely small, but this comparison was not made directly in the current analyses.

In analyses examining the effects of county expenditures on post-maltreatment competence for each program separately, similar results were found with respect to both More at Four and Smart Start. The only distinct finding was that participants living in counties with large expenditures on More at Four had somewhat greater odds of being competent following maltreatment than participants living in counties with no expenditures on More at Four (OR = 1.238, 95% CI [1.096, 1.399]). The other two comparisons (moderately large compared to no

expenditures and small compared to no expenditures) were not significantly different. The overall effect of More at Four was quite small, $F(3, 290) = 15.95$, $p < .0001$.

Third, multilevel logistic regression analyses were used to determine whether relative levels of expenditures on Smart Start and More at Four predicted participants' consistent competence across all eligible years of education data, including both pre- and post-maltreatment years. Participants who were categorized in the "always competent" subgroup were deemed to be consistently competent across both time periods for purposes of this set of analyses, and participants classified as "never competent" were classified as not competent in either time period. "Declined" indicated that the participant was competent prior to maltreatment and not competent after maltreatment. "Improved" indicated that the participant was not competent prior to maltreatment and competent after maltreatment.

Multilevel logistic regression was used to test the overall effect of relative levels of expenditures on Smart Start and More at Four on participants' consistent competence across all domains and across all years of eligible data. After adjusting for the full set of covariates, Smart Start and More at Four expenditures predicted consistently competent functioning across time, $F(2, 196) = 82.610$, $p < .0001$. Participants who lived in counties with relatively large expenditures on Smart Start and More at Four had twice the odds of being consistently competent than participants who lived in counties with small expenditures on Smart Start and More at Four ($OR = 2.015$, 95% CI [1.808, 2.246]), and participants who lived in counties with moderate levels of expenditures had 1.5 times the odds of being competent across all included years than participants who lived in counties with small expenditures ($OR = 1.535$, 95% CI [1.406, 1.677]).

In analyses examining the effects of expenditures on More at Four and Smart Start separately, positive effects on consistent competence across all available years of education data were found with respect to both More at Four and Smart Start. Relative levels of More at Four

expenditures within participants' county when they were age 4 predicted consistent competence across years after controlling for the effects of the covariates, $F(3, 290) = 191.350$, $p < .0001$. Participants who lived in counties with large expenditures on More at Four had four times the odds of being consistently competent than participants who lived in counties with no expenditures on More at Four (*OR* = 3.993, 95% CI [3.530, 4.518]), and participants who lived in counties with moderately large levels of More at Four expenditures had 2.5 times the odds of being competent across all eligible years than participants who lived in counties with no More at Four expenditures (*OR* = 2.530, 95% CI [2.258, 2.835]). Participants living in counties with moderate levels of expenditures had 1.5 times the odds of being consistently competent than participants living in counties with no More at Four expenditures (*OR* = 1.576, 95% CI [1.377, 1.805]).

After adjusting for covariates, levels of expenditures on Smart Start within participants' county when they were age 0-5 also predicted consistent competence across all pre- and post-maltreatment years, $F(3, 290) = 52.830$, $p < .0001$. Participants whose counties were allocated large amounts for the Smart Start program when participants were age 0-5 had somewhat greater odds of being consistently competent than participants whose counties were allocated small amounts for Smart Start (*OR* = 1.346, 95% CI [1.146, 1.580]). Participants who lived in counties with relatively large or moderate levels of Smart Start expenditures had twice the odds of being consistently competent across time than participants who lived in counties with small levels of Smart Start expenditures (*OR* = 1.914, 95% CI [1.641, 2.232], and *OR* = 2.062, 95% CI [1.782, 2.386], respectively).

4. Discussion

Using a large statewide sample of children, the current study produced estimates of consistently competent functioning within domains of academic performance, special education, and behavioral functioning across time of 18% for maltreated children and 35% for nonmaltreated children. Differences in competence between maltreated and nonmaltreated children were significant, with nonmaltreated children having about twice the odds of being competent across time as maltreated children. The study analyzed effects of child welfare policy change in the state and found that maltreated children were more likely to be consistently competent if the county investigating their reported maltreatment already had adopted the Multiple Response System prior to their maltreatment. Higher levels of per capita government expenditures on early childhood programs within children's counties also were found to promote competent outcomes among both maltreated and nonmaltreated children similarly.

The current study contributes to our understanding of the outcome of positive adaptation despite adverse childhood experiences like child maltreatment, referred to as resilience. Important outstanding questions about resilience among maltreated children that were addressed in the current study involve reliable measurement of resilient outcomes using large samples, longitudinal data, and multiple domains of functioning in order to calculate prevalence rates and compare characteristics of resilient and non-resilient children. This study also provides greater knowledge of community-level factors that can promote resilient outcomes, specifically, state policies and programs aimed at improving the child welfare system and enhancing child development and early learning. Another important contribution was the use of a large, statewide comparison sample of nonmaltreated children in some analyses to investigate rates of consistent competence compared to maltreated children, differences in characteristics of maltreated and

nonmaltreated children, and any differential impacts of expenditures on early childhood programs for maltreated and nonmaltreated children.

4.1 Prevalence of Resilience and Patterns of Competence

The prevalence of resilience, defined as post-maltreatment functioning at a minimally competent level in every domain at every time point, was 18%. The rate of resilience described in this study suggests a slightly more favorable outlook for the long-term developmental outcomes of maltreated children than typically has been found in the literature. However, the prevalence of resilience was low, with more than 4 out of 5 maltreated children failing to meet expectations at school consistently over time. Further, the fact that nonmaltreated children demonstrated twice the odds of being consistently competent as maltreated children after controlling for other risk factors for poor academic outcomes, such as low socioeconomic status and minority race/ethnicity status, indicates a compelling need for targeted intervention and support in the academic setting for these particularly high-risk children.

One potential challenge to the representativeness of the current study might be that the criteria for resilience in this study were relatively easy to satisfy. These criteria were selected as the most fundamental indicators of whether or not a child is meeting expectations in the school environment as measured by the state. Though a higher threshold of functioning could have been used in the present study, there is real-world significance to achieving a passing score on standardized testing, being promoted to the next grade, not requiring expensive special education services, and not engaging in problem behaviors that result in formal discipline. These outcomes all are readily measurable indicators of competent functioning that should characterize the average student. These criteria also involve generally normative criteria that apply to all students in the state, in contrast to studies that have used measures of competence based on functioning relative to that of other children within the specific study sample, some of which samples

included only maltreated or other high-risk children. The current study has the benefit of a large, statewide sample of nonmaltreated children to provide a normative view of competence within the public school setting among all students born and educated in the state.

Further evidence that the criteria for resilience were appropriate is reflected in the fact that the prevalence of consistent competence for nonmaltreated children in the study was 35%, and only about half of the nonmaltreated children were found to be competent in at least 75% of available years. In light of the fact that just over a third of nonmaltreated children met the standards for competence consistently across all years and domains, the criteria do not appear to be easy for the average student to meet. This relatively low rate of competence even among nonmaltreated children indicates that the criteria are fairly rigorous. Conversely, the fact that about half of nonmaltreated children met the standards for competence in most years and almost three-quarters of them were competent least half of the available years suggests that the standard for competence was not unreasonably high. When considering the appropriateness of the criteria used, it also is important to note that 94% of the nonmaltreated children and 81% of the maltreated children were competent across all domains in at least one year, and only a few were never competent in any domain (0.03% of nonmaltreated children and 0.5% of maltreated children). Taken together, these findings suggest that the criteria for competence have been set neither too low nor too high.

One limitation of the current study is that there is no direct measure of emotional functioning, particularly internalizing symptoms. The use of reported infractions as an indicator of behavioral functioning would be expected to capture some aspects of children's emotional disturbance, which can be reflected in problem behaviors at school for some children. Admittedly, it is likely that reports of behavioral infractions may involve more severe behaviors and may not include less severe behavior problems that could be measured more sensitively on questionnaires

of socioemotional functioning used in some previous studies. In addition, children's problem behaviors are not always observed at school, and some children whose emotional disturbance manifests primarily as internalizing symptoms do not exhibit problem behaviors that would give rise to behavioral infractions. Thus, it is possible that the lack of a direct measure of emotional functioning may inflate the prevalence of resilience found in the current study.

On the other hand, children's emotional functioning often impacts their academic performance and might be expected to decrease the likelihood of passing standardized testing or being promoted. Also, school accommodations and special education services sometimes are implemented for children with emotional disturbance, which would be included in an Individualized Education Plan and thus would be captured in the special education domain in the present study. Given the complex ways in which children's emotional functioning can influence their academic functioning across multiple domains, it seems probable that each of the indicators of resilience used in the present study reflects some aspects of emotional functioning even though more direct measures are not available in the state administrative data.

4.2 Effect of Differential Response Policies on Resilience

In addition to measuring the prevalence of resilience, this study examined the effect of the introduction of differential response policies in the state on the likelihood of resilient outcomes among maltreated children. Taking advantage of quasi-experimental differences in the timing of the implementation of the Multiple Response System across counties in the state, this study found that maltreated children who were reported for maltreatment in counties that had adopted differential response policies had twice the odds of resilience across subsequent years of school than maltreated children who were reported for maltreatment prior to the policy change in their county. To our knowledge, this study is the first to examine the effect of differential response policies on long-term developmental outcomes of maltreated children in a rigorous

design. The results of this study provide evidence that the ways in which child protective agencies interact and intervene with high-risk families, including referring families for preventive and early intervention services and adopting collaborative approaches that empower caregivers to participate in assessing and addressing their own family's needs, can have enduring impacts on the adaptive functioning of maltreated children.

Though this study did not examine the specific mechanisms through which differential response policies, including the use of family assessments, operated to promote resilient outcomes among maltreated children, several possible explanations exist. One possibility is that reduced rates of recurrence of maltreatment, which have been confirmed in recent evaluations of the Multiple Response System (e.g., Center for Child and Family Policy, 2009; Lawrence et al., 2011), mediate resilient outcomes. Though not included in the present analyses, a follow-on study could be conducted to measure frequencies of repeated reports of maltreatment to child protective services for maltreated children in the sample and test whether rates of recurrence differ across participants reported prior to and after the introduction of the Multiple Response System in their county. Assuming the predicted effect regarding the recurrence of maltreatment were found, mediation analyses then could test whether the reduction in maltreatment recurrence mediates the observed effect of the Multiple Response System on resilience.

Other possible explanations could be that increased service use and receipt of services earlier in the assessment process leads to lower parenting stress, more appropriate discipline strategies and effective parenting skills, better parent-child relationships, or improved coping for maltreated children. Additionally, providing preventive or early intervention services in less severe cases of maltreatment and reserving more intensive use of investigative resources for more severe cases of maltreatment arguably could have prevented some incidents of maltreatment altogether or at least may have resulted in reduced harm to children at risk. However, there is

limited reliable information about the severity of suspected or confirmed maltreatment as well as the type, duration, and quality of specific services provided to families reported for maltreatment, which makes exploring these potential mechanisms using the available data difficult.

4.3 Effect of Government Investments in Early Childhood Programs

Another community-level factor that was predicted to influence maltreated children's resilient outcomes was government investment in early childhood intervention, represented in the present study by county-level expenditures on two programs, Smart Start and More at Four. Differences in the date of adoption of these programs in counties across the state and variations in funding amounts across counties permitted use of a quasi-experimental design to investigate their effects on resilience. This study examined the full, community-wide effects of the programs as a whole and was not limited to participants known to have participated directly in each program, which information was not available for the present study.

Consistent with previous research finding positive impacts on children's academic outcomes from investments in early childhood programs, the current study found that total levels of expenditures on Smart Start and More at Four promoted adaptive outcomes among maltreated children and nonmaltreated children alike. These findings provide further support for the hypothesis that investments in early childhood promote adaptive functioning in disadvantaged children (e.g., Heckman & Masterov, 2007) and extend application of this hypothesis specifically to maltreated children.

4.4 Independent Effects of Differential Response Policies and Investments in Early Childhood Programs

Both the introduction of the Multiple Response System in North Carolina as well as higher levels of expenditures on Smart Start and More at Four were found to promote resilient outcomes among maltreated children. In light of these positive findings, supplemental analyses

examined whether differential response policies and government investments in early childhood programs exerted unique influences on resilience after controlling for the effects of the other factor. These analyses found largely independent effects of each of the two factors on resilience, suggesting that the mechanisms underlying each factor are distinct. As described above, the current study did not explore the specific mechanisms explaining how each factor operates upon maltreated children's academic outcomes.

4.5 Limitations

In describing the results presented in the current study, several inferences have been made regarding the effects of maltreatment on competence. Typically, these effects have been evidenced by group differences in competent outcomes according to participants' maltreatment status. It is important to note that these analyses have focused on the post-maltreatment functioning of maltreated children and do not control for maltreated children's functioning prior to experiences of maltreatment. A plausible alternative explanation for the differences in adaptive functioning between nonmaltreated and maltreated children found in the current study is that children's pre-maltreatment functioning influences their likelihood of being maltreated as well as their post-maltreatment functioning. The current analyses cannot adequately refute this possibility.

A possible solution to these limitations of the current findings would be to conduct additional analyses that include measures of pre-maltreatment functioning for those maltreated children who have available years of pre-maltreatment education data. Comparable measures of "early" functioning over the same academic periods represented in the pre-maltreatment functioning measures could be constructed for children in the comparison group. These measures then could be used as control variables in analyses comparing post-maltreatment functioning for maltreated children to an equivalent measure for "later" functioning created for nonmaltreated children. Any group differences in post-maltreatment competence that emerge after controlling

for children's pre-maltreatment functioning arguably would reflect the effects of maltreatment. This approach would provide a strong challenge to the argument that children's pre-maltreatment functioning, and not maltreatment, accounts for their post-maltreatment functioning.

4.6 Conclusion

Child maltreatment is a serious public health problem that has long-lasting psychological, social, physical, and economic impacts on its victims. Notwithstanding these grave consequences for many victims of maltreatment, some children demonstrate positive adaptation despite the extreme adversity they faced in childhood. The heterogeneity in outcomes among maltreated children poses many interesting questions for researchers, including: What proportion of maltreated children demonstrate resilience and in which domains of functioning? Is resilience stable over time and across domains? What factors appear to promote resilience in maltreated children? Through what processes do these factors have their effects on resilient development? These questions have not yet been addressed adequately and their answers have significant implications for prevention and intervention policy and practice.

The current study sought to inform child maltreatment research, policy, and practice, including advancing empirical knowledge about the prevalence, patterns, and predictors of resilient response to maltreatment. Maltreatment research to date has focused significant attention on identifying risk factors for maltreatment and predictors of negative outcomes, but substantially less effort has been devoted to identifying factors that foster positive adaptation among high-risk children. Greater understanding of factors and processes that promote resilient responses to maltreatment can meaningfully inform the contents, methods, and timing of prevention and intervention efforts targeting maltreated children or children at risk for maltreatment.

Specifically, information about the prevalence of resilience following maltreatment not only can clarify how common or rare resilient outcomes may be, but also serves as a broad

measure of how well society has responded to the long-term needs of maltreated children. In addition, patterns of instability in resilient outcomes can elucidate critical points of intervention for maltreated children even after involvement with child protection agencies typically ends. Also, greater knowledge of factors that promote resilience and processes underlying resilient development following maltreatment can help identify targets for building competencies in areas shown to support resilient development as well as specify mechanisms through which resilience may be fostered. Furthermore, given limited resources for universal and primary prevention programs, it may be useful to identify external factors that have a large impact on maltreated children, in order to inform policies and practices that may be especially effective for targeting adaptive outcomes in this high-risk group. In particular, state policies regarding child protection's response to families reported for maltreatment as well as early childhood programs intended to improve developmental outcomes for at-risk children may hold promise for promoting adaptive functioning among maltreated children.

The results of the current study contribute to greater understanding of resilience among maltreated children in multiple ways. It is the largest study of its kind that examined resilience as a multidimensional construct in a longitudinal design encompassing many years of data, allowing a more reliable estimate of the prevalence of resilience than previously available. The large sample size of maltreated children allowed closer examination of factors that contribute to resilience than has been possible in most previous studies, given the relatively low base rate of resilience following experiences of maltreatment in the population.

This study provided evidence of long-term benefits for maltreated children as a result of differential response policies in child welfare agencies, which include providing needs-based assessments for at-risk families when possible, fostering collaborative partnerships with parents to better assess families' many needs, and providing links to services early in the assessment

process to promote more effective parenting and better family functioning. The importance of investments in early childhood for promoting later adaptive functioning in the academic environment among maltreated children also was supported, particularly with respect to programs focused on providing high-quality early education and improving students' readiness to learn prior to entering school. Differential response and investments in early childhood were found to have independent effects on resilience, which suggests that both types of government policies and interventions make unique and meaningful contributions to positive adaptation among maltreated children.

References

Arnow, B. A. (2004). Relationships between childhood maltreatment, adult health and psychiatric outcomes, and medical utilization. *Journal of Clinical Psychiatry, 65*(Suppl. 12), 10-15.

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science, 333*, 975-978.

Barth, R. P., Scarborough, A., Lloyd, E. C., Losby, J., Casanueva, C., & Mann, T. (2007). *Developmental status and early intervention service needs of maltreated children*. Washington, DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.

Belsky, J. (1993). Etiology of child maltreatment: A developmental-ecological analysis. *Psychological Bulletin, 114*, 413-434.

Bolger, K. E., & Patterson, C. J. (2001). Pathways for child maltreatment to internalizing problems: Perceptions of control as mediators and moderators. *Development and Psychopathology, 13*, 913-940.

Bolger, K. E., & Patterson, C. J. (2003). Sequelae of child maltreatment: Vulnerability and resilience. In S. Luthar (Ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 156-181). New York, NY: Cambridge University Press.

Bolger, K. E., Patterson, C. J., & Kupersmidt, J. B. (1998). Peer relationships and self-esteem among children who have been maltreated. *Child Development, 69*, 1171-1197.

Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., & Chen, F. (2007). Latent variable models under misspecification: Two-stage least squares (2SLS) and maximum likelihood (ML) estimators. *Sociological Methods & Research, 36*, 48-86.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1986). Ecology of the family as a context to human development: Research perspectives. *Developmental Psychology, 22*, 723-742.

Bryant, D., Bernier, K., Peisner-Feinberg, E., & Maxwell, K. (2002). *Smart Start and child care in North Carolina: Effects on quality and changes over time.* Chapel Hill: University of North Carolina, FPG Child Development Institute.

Bryant, D., Maxwell, K., Taylor, K., Poe, M., Peisner-Feinberg, E., & Bernier, K. (2003). *Smart Start and preschool child care quality in North Carolina: Change over time and relation to children's readiness*. Chapel Hill, NC: University of North Carolina, FPG Child Development Institute.

CAPTA, The Keeping Children and Families Safe Act of 2003, 42 U.S.C. § 5106.

Center for Child and Family Policy. (2009, June). *Multiple Response System (MRS) evaluation report to the North Carolina Division of Social Services (NCDSS)*. Durham, NC: Duke University, Center for Child and Family Policy.

Center for Child and Family Policy. (2006, June). *Multiple Response System (MRS) evaluation report to the North Carolina Division of Social Services (NCDSS)*. Durham, NC: Duke University, Center for Child and Family Policy.

Center for Child and Family Policy. (2004, April). *Multiple Response System (MRS) evaluation report to the North Carolina Division of Social Services (NCDSS)*. Durham, NC: Duke University, Center for Child and Family Policy.

Cicchetti, D., & Blender, J. A. (2006). A multiple-levels-of-analysis perspective on resilience: Implications for the developing brain, neural plasticity, and preventive interventions. *Annals of the New York Academy of Sciences, 1094*, 248-258.

Cicchetti, D., & Lynch, M. (1993). Toward an ecological/transactional model of community violence and child maltreatment: Consequences for children's development. *Psychiatry, 56*, 96-118.

Cicchetti, D., & Manly, J. T. (Eds.). (2001). Operationalizing child maltreatment: Developmental processes and outcomes [Special Issue]. *Development and Psychopathology, 13*, 755-1048.

Cicchetti, D., & Rogosch, F. A. (1997). The role of self-organization in the promotion of resilience in maltreated children. *Development and Psychopathology, 9*, 797-815.

Cicchetti, D., & Toth, S. L. (2006). A developmental psychopathology perspective on preventive interventions with high risk children and families. In A. Renninger & I. Sigel (Eds.), *Handbook of Child Psychology* (6th ed., pp. 497-547). New York, NY: John Wiley & Sons, Inc.

Cicchetti, D., & Valentino, K. (2006). An ecological-transactional perspective on child maltreatment: Failure of the average expectable environment and its influence on child development. In D. Cicchetti and D. J. Cohen (Eds.), *Developmental psychopathology* (2nd ed., Vol. 3, pp. 129-201). New York, NY: Wiley.

Colman, R. A., & Widom, C. S. (2004). Childhood abuse and neglect and adult intimate relationships: A prospective study. *Child Abuse & Neglect, 28*, 1133-1151.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives, 15*, 213-218.

Daro, D., & Dodge, K. A. (2009). Creating community responsibility for child protection: Possibilities and challenges. *The Future of Children, 19*, 67-93.

Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). Mechanisms in the cycle of violence. *Science, 250*, 1678-1683.

Drake, B. (1996). Unraveling "unsubstantiated." *Child Maltreatment, 1*, 261-271.

Drake, B., Jonson-Reid, M., Way, I., & Chung, S. (2003). Substantiation and recidivism. *Child Maltreatment, 8*, 248-260.

Egeland, B., Carlson, E., & Sroufe, L. A. (1993). Resilience as process. *Development and Psychopathology, 5*, 517-528.

Egeland, B., Sroufe, L. A., & Erickson, M. (1983). The developmental consequences of different patterns of maltreatment. *Child Abuse and Neglect, 7*, 459-470.

Ewigman, B., Kivlahan, C., & Land, G. (1993). The Missouri Child Fatality Study: Underreporting of maltreatment fatalities among children younger than five years of age, 1983 through 1986. *Pediatrics, 91*, 330-337.

Fang, X., Brown, D. S., Florence, C. S., & Mercy, J. A. (2012). The economic burden of child maltreatment in the United States and implications for prevention. *Child Abuse & Neglect, 36*, 156-165.

Farber, E. A., & Egeland, B. (1987). Invulnerability among abused and neglected children. In E. J. Anthony and B. Cohler (Eds.), *The invulnerable child* (pp. 253-288). New York: Guilford Press.

Finkelhor, D., Turner, H. A., & Hamby, S. L. (2005). The victimization of children and youth: A comprehensive, national survey. *Child Maltreatment*, *10*(1), 5-25.

Flores, E., Cicchetti, D., & Rogosch, F. A. (2005). Predictors of resilience in maltreated and nonmaltreated Latino children. *Developmental Psychology, 41*, 338-351.

Fraser, M. W. (2004). The ecology of childhood: A multisystems perspective. In M. W. Fraser (Ed.), *Risk and resilience in childhood: An ecological perspective* (2nd ed., pp. 1-12). Washington, DC: NASW Press.

Haskett, M. E., Nears, K., Ward, C. S., & McPherson, A. V. (2006). Diversity in adjustment of maltreated children: Factors associated with resilient functioning. *Clinical Psychology Review, 26*, 796-812.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900-1902.

Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Review of Agricultural Economics, 29*, 446-493.

Herrenkohl, E. C., Herrenkohl, R. C., & Egolf, B. (1994). Resilient early school-age-children from maltreating homes: Outcomes in late adolescence. *American Journal of Orthopsychiatry, 64*, 301-309.

Hussey, J. M., Marshall, J. M., English, D. J., Knight, E. D., Lau, A. S., Dubowitz, H., & Kotch, J. B. (2005). Defining maltreatment according to substantiation: Distinction without a difference? *Child Abuse and Neglect, 29*, 479-492.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Jackson, S., & Martin, P. Y. (1998). Surviving the care system: Education and resilience. *Journal of Adolescence, 21*, 569-583.

Jaffee, S. R., Caspi, A., Moffitt, T. E., Polo-Tomas, M., & Taylor, A. (2007). Individual, family, and neighborhood factors distinguish resilient from non-resilient maltreated children: A cumulative stressors model. *Child Abuse and Neglect, 31*, 231-253.

Jaffee, S. R., & Gallop, R. (2007). Social, emotional, and academic competence among children who have had contact with child protective services: Prevalence and stability estimates. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 757-765.

Kaplan, C., & Merkel-Holguin, L. (2008). Another look at the national study on differential response in child welfare. *Protecting Children*, *23*, 5-21.

Kaplan, S. J., Pelcovitz, D., & Labruna, V. (1999). Child and adolescent abuse and neglect research: A review of the part 10 years. Part I: Physical and emotional abuse and neglect. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1214-1222.

Kaufman, J., Cook, A., Arny, L., Jones, B., & Pittinsky, T. (1994). Problems defining resiliency: Illustrations from the study of maltreated children. *Development and Psychopathology, 6*, 215-229.

Kaufman, J., & Zigler, E. (1987). Do abused children become abusive parents? *American Journal of Orthopsychiatry, 57*, 186-198.

Kim, J., & Cicchetti, D. (2004). A longitudinal study of child maltreatment, mother-child relationship quality and maladjustment: The role of self-esteem and social competence. *Journal of Abnormal Child Psychology, 32*, 341–354.

Kohl, P. L., Jonson-Reid, M., & Drake, B. (2009). Time to leave substantiation behind: Findings from a national probability study. *Child Maltreatment, 14*, 17–26.

Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2012). *From birth to school: Early childhood programs and third grade outcomes in North Carolina.* Manuscript in preparation.

Lawrence, C. N., Rosanbalm, K. D., & Dodge, K. A. (2011). Evaluation of policy change in North Carolina's child welfare system. *Children and Youth Services Review, 33*, 2355-2365.

Leiter, J., Myers, K. A., & Zingraff, M. T. (1994). Substantiated and unsubstantiated cases of child maltreatment: Do their consequences differ? *Social Work Research, 18*(2), 67-82.

Lerner, R. M. (1998). Theories of human development: Contemporary perspectives. In W. Damon & R. Lerner (Eds.), *Handbook of child psychology Vol. 1: Theoretical models of human development* (5th ed., pp. 1-24). New York, NY: Cambridge University Press.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for Mixed Models* (2nd ed.). Cary, NC: SAS Institute, Inc.

Loman, L. A. (2005, May). Differential response improves traditional investigations: Criminal arrests for severe physical and sexual abuse. St. Louis, MO: Institute of Applied Research.

Loman, A. L., & Siegel, G. L. (2004). *Differential response in Missouri after five years: Final report.* St. Louis, MO: Institute of Applied Research.

Loman, A. L., & Siegel, G. L. (2004). *Minnesota alternative response evaluation: Final report*. St. Louis, MO: Institute of Applied Research.

Love, J. M., Kisker, E. E., Ross, C. M., Schochet, P. Z., Brooks-Gunn, J., Paulsell, D., . . . Brady-Smith, C. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start, Vol. 1: Final Technical Report*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Luthar, S. S. (1991). Vulnerability and resilience: A study of high-risk adolescents. *Child Development, 62*, 600-616.

Luthar, S. S., & Cicchetti, D. (2000). The construct of resilience: Implications for interventions and social policies. *Development and Psychopathology, 12*, 857-885.

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development, 71*, 543-562.

Luthar, S. S., & Zigler, E. (1991). Vulnerability and competence: A review of research on resilience in childhood. *American Journal of Orthopsychiatry, 61*, 6-22.

Masten, A. S. (2006). Developmental psychopathology: Pathways to the future. *International Journal of Behavioral Development, 30*, 47-54.

Masten, A. S., & Coatesworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist, 53*, 205-220.

Masten, A. S., & Wright, M. O. D. (1998). Cumulative risk and protection models of child maltreatment. *Journal of Aggression, Maltreatment & Trauma, 2*(1), 7-30.

McGloin, J. M., & Widom, C. S. (2001). Resilience among abused and neglected children grown up. *Development and Psychopathology, 13*, 1021-1038.

Merkel-Holguin, L., Kaplan, C., & Kwak, A. (2006). *National study on differential response in child welfare*. Washington, DC: American Humane Association and Child Welfare League of America.

Newman, D., & Newman, I. (2012). Multilevel modeling: Clarifying issues of concern. *Multiple Linear Regression Viewpoints, 38*, 26-33.

North Carolina Department of Health and Human Services. (2003). North Carolina embarks on major reform of its child welfare system. *Fostering Perspectives, 7*(2). Raleigh, NC: North Carolina Department of Health and Human Services, Division of Social Services.

Olds, D., Eckenrode, J., Henderson, C., Kitzman, H., Powers, J., Cole, R., . . . Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association, 278*, 637-643.

Peisner-Feinberg, E. S., & Schaaf, J. M. (2010). *Long-term effects of the North Carolina More at Four Pre-kindergarten Program: Children's reading and math skills at third grade.* Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.

Perkins, D., & Jones, K. (2004). Risk behaviors and resiliency within physically abused adolescents. *Child Abuse and Neglect, 28*, 547-563.

Putnam, F. W. (2003). Ten-year research update review: Child sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(3), 269-278.

Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science, 4*, 2-14.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Sagy, S., & Dotan, N. (2001). Coping resources of maltreated children in the family: A salutogenic approach. *Child Abuse and Neglect, 25*, 1463-1480.

Sameroff, A. J. (2000). Developmental systems and psychopathology. *Development and Psychopathology, 12*, 297-312.

SAS Institute Inc. (2011). SAS® (Version 9.3) [Computer software]. Cary, NC: Author.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* London, England: Chapman & Hall.

Schwartz, D., Dodge, K. A., Pettit, G. S., Bates, J. E., & The Conduct Problems Prevention Research Group. (2000). Friendship as a moderating factor in the pathway between early harsh home environment and later victimization in the peer group. *Developmental Psychology, 36*, 646-662.

Schweinhart, L., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry preschool study through age 40*. Ypsilanti, MI: High Scope Press.

Sedlak, A. J., Mettenburg, J., Basena, M., Petta, I., McPherson, K., Greene, A., & Li, S. (2010). *Fourth National Incidence Study of Child Abuse and Neglect (NIS–4): Report to Congress, Executive Summary*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Siegel, G. L., & Loman, T. (2006). *Extended follow-up study of Minnesota's Family Assessment Response: Final report*. St. Louis, MO: Institute of Applied Research.

Snijders, T. A. B., & Bosker, B. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: Sage Publications.

Stagner, M., & Lansing, J. (2009). Progress toward a prevention perspective. *The Future of Children, 19*, 19-38.

Thornberry, T. P., Ireland, T. O., & Smith, C. A. (2001). The importance of timing: The varying impact of childhood and adolescent maltreatment on multiple problem outcomes. *Development and Psychopathology, 13*, 957-979.

Toth, S. L., & Cicchetti, D. (2006). Promises and possibilities: The application of research in the area of child maltreatment to policies and practices. *Journal of Social Issues, 62*, 863-880.

U.S. Department of Health and Human Services. (2006). *Research to practice: Preliminary findings from the Early Head Start prekindergarten follow-up, Early Head Start Research and Evaluation Project*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

U.S. Department of Health and Human Services. (2008). *Child Maltreatment 2006*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

U.S. Department of Health and Human Services. (2012). *Child Maltreatment 2011*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Vogel, C. A., Xue, Y., Moiduddin, E. M., Carlson, B. L., & Kisker, E. E. (2010). *Early Head Start children in grade 5: Long-term follow-up of the Early Head Start Research and Evaluation Study sample*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Waldfogel, J. (2009). Differential response. In K. A. Dodge & D. L. Coleman (Eds.), *Preventing child maltreatment: Community approaches*. New York, NY: Guilford Press.

Walsh, W. A., Dawson, J., & Mattingly, M. J. (2010). How are we measuring resilience following childhood maltreatment? Is the research adequate and consistent? What is the impact on research, practice, and policy? *Trauma, Violence, & Abuse, 11*, 27-41.

Wang, C-T., & Holton, J. (2007). *Total estimated cost of child abuse and neglect in the United States*. Prevent Child Abuse America.

Widom, C. S. (1999). Posttraumatic stress disorder in abused and neglected children grown up. *American Jorunal of Psychiatry, 156*(8), 1223-1229.

Wilson, H. W., & Widom, C. S. (2008). An examination of risky sexual behavior and HIV in victims of child abuse and neglect: a 30-year follow-up. *Health Psychology, 27*, 149-58.

Biography

Sandra Y. Nay McCourt was born in Bloomington, IN in 1972. She attended college at DePauw University, earning a Bachelor of Arts in Classical Greek and Political Science in 1994. She then attended law school at the University of Michigan Law School and was an associate editor of the Michigan Journal of International Law. Sandra earned a Juris Doctor in 1996. Sandra subsequently became interested in a career in policy-oriented social science research and studied psychology at New York University and Columbia University in New York, NY, eventually earning a Certificate in Psychology from Columbia in 2005. Sandra enrolled in the Ph.D. program in clinical psychology at Duke University, earning a Master of Arts and a Certificate in Education Policy Research in 2009. During the 2012-2013 academic year, Sandra has been completing her clinical psychology internship at the Indiana University School of Medicine in Indianapolis, IN.

During the course of her doctoral studies at Duke, Sandra was awarded the James B. Duke Fellowship, Phillip Jackson Baugh Fellowship, and Summer Research Fellowship through Duke University, the Spencer Foundation Education Science and Policy Fellowship through the Center for Child and Family Policy at Duke University, the National Science Foundation Graduate Student Fellowship, the National Institute for Child Health and Human Development Predoctoral Fellowship through the Center for Developmental Science at the University of North Carolina, the Elizabeth Munsterberg Koppitz Graduate Student Fellowship in Child Psychology through the American Psychological Foundation, and the Doris Duke Fellowship for the Promotion of Child Well-Being through the Doris Duke Charitable Foundation and Chapin Hall at the University of Chicago. Sandra also was inducted into the Psi Chi International Honor Society in Psychology and participated as a fellow in the Preparing Future Faculty program at Duke.